

Contributions à l'interprétation de mesures non additives et à l'identification de modèles décisionnels fondés sur l'intégrale de Choquet

MÉMOIRE

présenté et soutenu publiquement le 21 novembre 2006

pour l'obtention de l'

Habilitation de l'Université de Nantes

Spécialité Informatique

par

Ivan Kojadinovic

Composition du jury

- Président :* Michel Grabisch, Professeur, LIP6, Université Paris I
- Rapporteurs :* Denis Bouyssou, Directeur de Recherche CNRS, LAMSADE
Thierry Dencœux, Professeur, HEUDIASYC, UTC
Didier Dubois, Directeur de Recherche CNRS, IRIT
- Examineurs :* Xavier Gandibleux, Professeur, LINA, Université de Nantes
Christophe Labreuche, Chercheur, THALES Research and Technology
- Coordinatrice :* Pascale Kuntz, Professeur, LINA, Université de Nantes

Mis en page avec la classe thloria.

Remerciements

Je souhaite exprimer ma profonde gratitude à Denis Bouyssou, Didier Dubois et Thierry Denœux pour avoir accepté de rapporter sur ces travaux. Un très, TRÈS grand merci également à Xavier Gandibleux, Michel Grabisch et Christophe Labreuche pour leur participation au jury.

Les travaux présentés dans ce document sont le résultat de fructueuses collaborations avec Jean-Luc Marichal, Katsushige Fujimoto, Michel Grabisch et Patrick Meyer. Je les remercie tous chaleureusement.

Enfin, un grand merci à mes collègues enseignants-chercheurs de l'École polytechnique de l'université de Nantes pour leur bonne humeur et leur soutien et, en particulier, à Pascale Kuntz, pour avoir encouragé mes activités de recherche et pour avoir accepté de coordonner cette habilitation.

À Chip, Mini-Bat, Dodo et tous les autres.

Table des matières

Préambule	1
Introduction	3
Chapitre 1 Jeux, capacités, intégrale de Choquet et généralisations	7
1.1 Introduction	8
1.2 Fonctions d'ensemble, jeux et capacités	8
1.2.1 Notations	8
1.2.2 Fonctions d'ensemble	9
1.2.3 Représentations équivalentes d'une fonction d'ensemble	9
1.2.4 Jeux et capacités	10
1.2.5 Fonctions d'ensemble k -additives	11
1.2.6 Le treillis $(\mathcal{P}(N), \subseteq)$	12
1.3 L'intégrale de Choquet discrète en tant qu'opérateur d'agrégation	13
1.3.1 Définition	14
1.3.2 Caractérisation axiomatique en tant qu'opérateur d'agrégation	14
1.3.3 Interprétation géométrique	15
1.3.4 Lien avec l'extension de Lovász	16
1.3.5 Premières extensions de l'intégrale de Choquet	16
1.4 Généralisations	17
1.4.1 Jeux bi-coopératifs et bi-capacités	17
1.4.2 L'inf-demi-treillis $(\mathcal{Q}(N), \subseteq)$	18
1.4.3 Transformée de Möbius sur $(\mathcal{Q}(N), \subseteq)$	20
1.4.4 L'intégrale de Choquet par rapport à un jeu bi-coopératif	21
Chapitre 2 Interprétation de mesures non additives	23
2.1 Introduction	24
2.2 Notion de valeur	25
2.2.1 Valeur de Shapley	25

2.2.2	Valeurs généralisées	27
2.2.3	Indices d'importance dans le cas bipolaire	27
2.2.4	Valeur de Shapley par rapport à un jeu bi-coopératif	31
2.3	Indices d'interaction	31
2.3.1	Notion d'interaction	31
2.3.2	Indices d'interaction de Shapley par rapport à un jeu coopératif	32
2.3.3	Extension au cas bipolaire	33
2.4	Mesures d'uniformité	34
2.4.1	Notion d'entropie	34
2.4.2	Extension de l'entropie de Shannon aux capacités normalisées	35
2.4.3	Variance d'une capacité normalisée	37
2.4.4	Interprétation dans le contexte de l'agrégation par intégrale de Choquet	38
2.4.5	Extension de l'entropie de Shannon aux bi-capacités normalisées	39
2.5	Loi et moments de l'intégrale de Choquet dans le cas uniforme	41
2.5.1	Définitions : fonctions puissances tronquées et différences divisées . . .	41
2.5.2	Principaux résultats	41
2.5.3	Calcul de différences divisées de fonctions puissances tronquées	42
2.5.4	Exemple	43
 Chapitre 3 Identification de modèles décisionnels fondés sur l'intégrale de Choquet		45
3.1	Introduction	46
3.2	Théorie de l'utilité multiattribut fondée sur l'intégrale de Choquet	47
3.2.1	Contexte théorique	47
3.2.2	Formulation du problème d'identification de capacités	48
3.3	Principales méthodes d'identification de capacités	50
3.3.1	Approches fondées sur les moindres carrés	50
3.3.2	Approche fondée sur la programmation linéaire	51
3.3.3	Méthodes du minimum de variance et du minimum de distance	52
3.3.4	Moindres carrés généralisés	53
3.3.5	Implémentation des méthodes d'identification	54
3.4	Exemples d'utilisation des méthodes d'identification	55
3.4.1	Description du problème	55
3.4.2	Approches fondées sur les moindres carrés	56
3.4.3	Approches linéaire, du minimum de variance et du minimum de distance	58
3.4.4	Contraintes supplémentaires sur la valeur de Shapley	60
3.4.5	Contraintes supplémentaires sur les indices d'interaction	62

3.4.6	Solution plus simple mais approchée	64
	Conclusion et perspectives	67
	Bibliographie	71
	Annexes	79
	Annexe A Articles de revues joints	81

Préambule

Ce document présente une synthèse des activités de recherche que j'ai menées depuis début 2002 sur les *mesures non additives*. J'ai choisi, par souci de clarté, de présenter en détails les résultats concernant l'*interprétation* et l'*identification* de mesures non additives obtenus dans le cadre de diverses collaborations et de ne décrire que succinctement les recherches connexes que je mène de façon moins soutenue sur la classification et la sélection de variables dans le contexte de l'analyse de données et de la modélisation statistique (cf. Annexes A et A).

Les travaux sur les mesures et intégrales non additives décrits dans ce document s'organisent selon deux axes : le premier axe porte sur l'*interprétation* de mesures non additives dans le contexte de la théorie des jeux coopératifs et de l'aide multicritère à la décision fondée sur les intégrales non additives ; le deuxième axe concerne l'*identification* de mesures non additives dans le contexte de l'agrégation par intégrale de Choquet. Ces deux directions de recherche sont fortement complémentaires dans la mesure où l'utilisation d'intégrales non additives dans un contexte décisionnel nécessite, dans un premier temps, l'identification d'une mesure non additive et, dans un deuxième temps, son interprétation afin de rendre le modèle de décision plus transparent pour le décideur.

Afin de rendre la lecture de ce document plus agréable, je l'ai ponctué de devises Shadoks particulièrement adaptées à la vie scientifique de l'enseignant-chercheur. Commençons par la devise ci-après :



Jacques Rouxel, *Les Shadoks*.

Elle vient rappeler, dans mon cas, la très forte inspiration que j'ai trouvée dans les travaux fondateurs de Michel Grabisch, Christophe Labreuche et Jean-Luc Marichal.

Notez qu'afin d'être en règle en termes de droits de reproduction, j'ai adressé une demande en ce sens au principal éditeur des *Shadoks* mais ma demande est restée sans réponse.

Introduction

La notion de *mesure* est un concept fondamental en mathématiques et repose sur la propriété d'*additivité* (voir p. ex. Rudin 1975). Cette propriété, bien que techniquement très commode, n'est pas appropriée dans certaines situations réelles dans lesquelles la quantité à modéliser n'est pas additive par essence. Afin de pallier le manque de flexibilité des mesures classiques, la propriété d'additivité a été abandonnée dans de nombreux domaines comme en théorie des fonctions de croyance (Dempster 1967, Shafer 1976, Smets & Kennes 1994, Klir & Yuan 1995), en théorie des possibilités (Dubois & Prade 1988), en théorie des jeux coopératifs (Curiel 1997, Peleg & Sudhölter 2003) ou en aide multicritère à la décision fondée sur les intégrales non additives (Grabisch & Labreuche 2004). Les fonctions d'ensemble à la base de ces théories satisfont des propriétés plus ou moins contraignantes et leur interprétation dépend du domaine considéré. L'ensemble de référence sous-jacent, noté N dans la suite, peut par exemple représenter (Grabisch 2005a) :

- l'ensemble des *états de la nature* : les sous-ensembles de N peuvent alors être interprétés comme des *événements* et les mesures non additives servent dans ce cas à modéliser l'*incertitude* relative au véritable état de la nature comme en théorie des fonctions de croyance ;
- un ensemble de *critères*, d'*attributs* ou de *capteurs* : dans ce contexte, les mesures non additives, appelées généralement *capacités* (Choquet 1953) ou *mesures floues* (Sugeno 1974), servent à modéliser l'*importance* des groupes de critères, d'attributs ou de capteurs comme en aide multicritère à la décision ou dans le domaine de la fusion de données fondées sur les intégrales non additives ;
- un ensemble de *votants* : le contexte est alors celui de la théorie des votes (voir p. ex. Shapley & Shubik 1954, Banzhaf 1965) ; les sous-ensembles de N sont appelés dans ce cas des *coalitions*, les mesures non additives servant à modéliser leur *pouvoir* ;
- un ensemble de *joueurs*, d'*agents*, d'*entreprises*, etc : les sous-ensembles de N sont alors également appelés des *coalitions* et les mesures non additives, appelées des *jeux*, servent à modéliser leur *gain* (souvent monétaire) ; le contexte est ici celui de la théorie des jeux coopératifs.

Valeur de Shapley et intégrale de Choquet

L'utilisation de fonctions d'ensemble dans le cadre des théories citées ci-dessus a mis en avant la nécessité de définir des concepts mathématiques permettant de *synthétiser* l'information qu'elles modélisent et de les *étendre* à des situations plus générales. Deux notions fondamentales répondant à ces deux préoccupations sont respectivement la *valeur de Shapley* (1953), initialement introduite dans le contexte de la théorie des jeux coopératifs, et l'*intégrale de Choquet*

(1953), définie comme une intégrale par rapport à une mesure monotone¹.

La valeur de Shapley d'une mesure non additive peut être regardée dans de nombreux domaines comme l'approximation additive la plus naturelle de cette fonction d'ensemble. Elle est connue sous le nom de *transformation pignistique* en théorie des fonctions de croyance (Smets & Kennes 1994, Smets 2005), sous le nom d'*indice d'importance de Shapley* en aide multicritère à la décision ou encore sous le nom d'*indice de pouvoir de Shapley & Shubik (1954)* en théorie des votes. Nous détaillerons les différentes interprétations de ce concept dans le Chapitre 2.

L'intégrale de Choquet par rapport à une mesure non additive sur N peut être regardée comme une extension de cette mesure à l'ensemble de sous-ensembles flous de N (Grabisch 2000). En théorie des jeux coopératifs, elle est connue sous le nom d'*extension de Lovász (1983)*. Dans le contexte de l'aide multicritère à la décision, l'intégrale de Choquet apparaît comme le moyen le plus simple pour étendre à des alternatives quelconques le raisonnement d'un décideur sur des alternatives *binaires* (Grabisch 2004, Grabisch & Labreuche 2004, Grabisch 2005b).

Généralisations des mesures non additives

De nombreuses généralisations des fonctions d'ensembles utilisées dans les domaines précédemment cités ont été récemment proposées afin de permettre une modélisation plus adéquate de la réalité. La théorie des jeux coopératifs s'est par exemple vue enrichie des notions de *jeu sur une géométrie convexe* (Bilbao 1998), de *jeu bi-coopératif* (Bilbao 2000) et de *jeu flou* (Aubin 1981). En théorie des votes, la notion de *jeu de vote ternaire* (Felsenthal & Machover 1997) est venue compléter l'existant afin de permettre la prise en compte de l'abstentionnisme. Dans le contexte de l'aide multicritère à la décision fondée sur les intégrales non additives, les *bi-capacités* sont venues étendre les capacités afin de permettre la prise en compte d'échelles d'évaluation *bipolaires* (Grabisch & Labreuche 2005a). Une synthèse des nombreuses extensions proposées dans la littérature peut être trouvée dans (Grabisch 2005a).

L'apparition de ces généralisations a naturellement conduit à la recherche d'extensions adéquates des outils fondamentaux que sont la valeur de Shapley et l'intégrale de Choquet. Un indice de pouvoir adapté aux jeux de vote ternaires a par exemple été défini par Felsenthal & Machover (1997). Grabisch & Labreuche (2005b) ont proposé une extension naturelle de l'intégrale de Choquet aux bi-capacités. Une extension raisonnable de l'indice d'importance de Shapley a été suggérée dans ce contexte par Kojadinovic (2006c). Labreuche & Grabisch (2006b) ont obtenu indépendamment une définition très similaire dans le cadre de la théorie des jeux bi-coopératifs.

Contenu et organisation de ce document

Dans le cadre de ce mémoire, nous nous intéressons aux mesures non additives et à certaines de leurs généralisations, principalement dans le contexte de la théorie des jeux coopératifs et de l'aide multicritère à la décision fondée sur les intégrales non additives. Le premier chapitre de ce document est consacré aux notions de *jeu* et de *capacité*, à la présentation de l'intégrale de Choquet en tant qu'opérateur d'agrégation ainsi qu'aux généralisations évoquées précédemment. Il s'agit essentiellement d'un chapitre de définitions. Les contributions aux domaines cités ci-dessus seront présentées dans le second et le troisième chapitre portant respectivement sur

¹La définition de l'intégrale de Choquet serait en fait due à Vitali (1925) et aurait été indépendamment redécouverte par Choquet (1953). Il est à noter que l'appellation "intégrale" fait débat chez les mathématiciens.

l'*interprétation* de mesures non additives et sur l'*identification* de modèles décisionnels fondés sur l'intégrale de Choquet.

La valeur de Shapley, bien qu'étant d'une importance fondamentale, ne permet pas de rendre compte de toute la complexité des phénomènes modélisés par une mesure non additive. Parmi les indices complémentaires qui ont été proposés dans la littérature pour permettre une interprétation plus complète d'une mesure non additive, les plus importants sont probablement les *indices d'interaction* (Murofushi & Soneda 1993, Grabisch 1997b) et les *mesures d'uniformité* (Marichal & Roubens 2000a, Kojadinovic 2006a). Ces concepts ont été récemment étendus aux bi-capacités dans (Grabisch & Labreuche 2005a, Kojadinovic 2006c, Kojadinovic & Marichal 2006, Labreuche & Grabisch 2006b) et seront étudiés dans le Chapitre 2. Nous y présenterons également les résultats obtenus dans (Marichal & Kojadinovic 2006) donnant la loi ainsi que les moments de l'intégrale de Choquet dans le cas uniforme.

L'intégrale de Choquet est à la base de nombreux modèles de décision dans l'incertain et le risque (voir p. ex. Schmeidler 1989, Tversky & Kahneman 1992, Chateauneuf 1994, Narukawa & Murofushi 2004). Elle peut être également vue comme un opérateur d'agrégation particulièrement flexible, ce qui suggère de l'utiliser pour la représentation numérique des préférences en aide multicritère à la décision (Grabisch 1995a, Marichal 2000a, Grabisch 2005b). Le troisième chapitre de ce mémoire est consacré à l'identification de modèles décisionnels fondés sur l'intégrale de Choquet. Dans le cadre d'un état de l'art des principales méthodes d'identification existant dans la littérature, nous présenterons les approches du *minimum de variance* et du *minimum de distance* proposées dans (Kojadinovic 2006a, Kojadinovic 2006b). Nous concluons le troisième chapitre par des exemples d'utilisation du *package* Kappalab (Grabisch, Kojadinovic & Meyer 2005) pour GNU R (R Development Core Team 2005) dans lequel ont été implémentées les principales méthodes d'identification de capacités.

Jeux, capacités, intégrale de Choquet et généralisations



Jacques Rouxel, *Les Shadoks*.

Sommaire

1.1	Introduction	8
1.2	Fonctions d'ensemble, jeux et capacités	8
1.2.1	Notations	8
1.2.2	Fonctions d'ensemble	9
1.2.3	Représentations équivalentes d'une fonction d'ensemble	9
1.2.4	Jeux et capacités	10
1.2.5	Fonctions d'ensemble k -additives	11
1.2.6	Le treillis $(\mathcal{P}(N), \subseteq)$	12
1.3	L'intégrale de Choquet discrète en tant qu'opérateur d'agrégation	13
1.3.1	Définition	14
1.3.2	Caractérisation axiomatique en tant qu'opérateur d'agrégation . . .	14
1.3.3	Interprétation géométrique	15

1.3.4	Lien avec l'extension de Lovász	16
1.3.5	Premières extensions de l'intégrale de Choquet	16
1.4	Généralisations	17
1.4.1	Jeux bi-coopératifs et bi-capacités	17
1.4.2	L'inf-demi-trellis $(\mathcal{Q}(N), \subseteq)$	18
1.4.3	Transformée de Möbius sur $(\mathcal{Q}(N), \subseteq)$	20
1.4.4	L'intégrale de Choquet par rapport à un jeu bi-coopératif	21

1.1 Introduction

Les mesures non additives sont des généralisations des mesures classiques permettant de modéliser des quantités qui ne sont pas *additives* par essence. Leur utilisation dans le cadre d'applications repose sur de nombreux outils mathématiques provenant de domaines aussi divers que la théorie des capacités, la combinatoire, la recherche opérationnelle ou la théorie des jeux coopératifs.

La première section de ce chapitre traite des notions fondamentales de *représentation équivalente* d'une fonction d'ensemble et de *k-additivité*, et introduit les concepts de *jeu* et de *capacité* dans le contexte de la théorie des jeux coopératifs et de l'aide multicritère à la décision respectivement. La seconde section est dévolue à l'*intégrale de Choquet*, qui peut être considérée comme l'intégrale la plus naturelle par rapport à une capacité, et, en particulier, à son utilisation en tant qu'opérateur d'agrégation. La dernière section présente les extensions des notions de jeu et de capacité récemment proposées dans la littérature ainsi qu'une généralisation naturelle de l'intégrale de Choquet permettant de prendre en compte, dans le contexte de l'agrégation, des évaluations définies sur une échelle *bipolaire*.

1.2 Fonctions d'ensemble, jeux et capacités

L'objectif de cette section est de définir les principaux concepts mathématiques gravitant autour des notions de *jeu* et de *capacité*.

1.2.1 Notations

Dans la suite de ce document, nous considérons un ensemble $N := \{1, \dots, n\}$ dont les éléments, en fonction du domaine considéré, représenteront des critères, des joueurs, des votants, des attributs, etc. L'ensemble des parties de N sera noté $\mathcal{P}(N)$. Le *vecteur caractéristique* d'un sous-ensemble $S \subseteq N$ sera noté $(1_S, 0_{N \setminus S})$. Il s'agit du vecteur de $\{0, 1\}^n$ dont la i -ème composante vaut 1 si $i \in S$, 0 sinon. Notons que, géométriquement, les vecteurs caractéristiques correspondent aux 2^n sommets de l'hypercube $[0, 1]^n$.

Afin d'éviter des notations trop lourdes, nous adoptons celles fréquemment utilisées dans la littérature (voir p. ex. Grabisch & Roubens 1999). Ainsi, nous omettrons souvent les accolades pour les singletons de N , p. ex., en écrivant i , $N \setminus i$ au lieu de $\{i\}$, $N \setminus \{i\}$. De même, pour les paires, nous écrirons ij au lieu de $\{i, j\}$, etc. De plus, le cardinal de sous-ensembles S, T, \dots , de N sera souvent noté par la lettre minuscule correspondante s, t, \dots .

1.2.2 Fonctions d'ensemble

Définition 1.2.1 Une fonction définie sur une famille de sous-ensembles d'un même ensemble est appelée fonction d'ensemble.

Soit μ une fonction d'ensemble définie sur $\mathcal{P}(N)$ à valeurs dans \mathbb{R} . La fonction μ est dite

- *monotone* lorsque, pour tous sous-ensembles A, B tels que $A \subseteq B$, $\mu(A) \leq \mu(B)$;
- *additive* lorsque, pour tous sous-ensembles A, B disjoints, $\mu(A \cup B) = \mu(A) + \mu(B)$;
- *cardinale* ou *symétrique* lorsque, pour tout $T \subseteq N$, $\mu(T)$ ne dépend que du cardinal de T .

Il existe alors $n+1$ réels $\mu_0, \mu_1, \dots, \mu_n$ tels que $\mu(T) = \mu_t$ pour tout $T \subseteq N$ tel que $|T| = t$.

1.2.3 Représentations équivalentes d'une fonction d'ensemble

Toute fonction d'ensemble $\mu : \mathcal{P}(N) \rightarrow \mathbb{R}$ est définie de façon unique par la donnée de la famille des 2^n nombres réels $(\mu(S))_{S \subseteq N}$. Si μ est additive, elle est complètement définie par ses valeurs sur les singletons de N , c.-à-d. par la famille $(\mu(i))_{i \in N}$.

Soit $\mathcal{T} : \mathbb{R}^{2^n} \rightarrow \mathbb{R}^{2^n}$ une transformation qui permet d'obtenir une autre famille de 2^n coefficients à partir de la famille $(\mu(S))_{S \subseteq N}$. Si la transformation \mathcal{T} est une bijection, la fonction d'ensemble μ est parfaitement définie par la donnée de la famille $\mathcal{T}((\mu(S))_{S \subseteq N})$. Dans ce cas, nous dirons que la fonction d'ensemble dont les coefficients sont donnés par $\mathcal{T}((\mu(S))_{S \subseteq N})$ est une *représentation équivalente* de μ (Grabisch 1997b, Grabisch 2000, Grabisch, Marichal & Roubens 2000).

La *représentation* ou *transformée de Möbius* d'une fonction d'ensemble (voir p. ex. Rota 1964, Grabisch, Marichal & Roubens 2000), appelée aussi *dividendes* en théorie des jeux coopératifs (Harsanyi 1963), est une représentation couramment utilisée en combinatoire, en aide multicritère à la décision fondée sur les intégrales non additives, en théorie des possibilités ou encore en théorie des fonctions de croyance. La représentation de Möbius d'une fonction d'ensemble μ sur N est définie par

$$m_\mu(S) := \sum_{T \subseteq S} (-1)^{s-t} \mu(T), \quad \forall S \subseteq N.$$

La transformation inverse, permettant d'obtenir μ à partir de m_μ , est connue sous le nom de *transformation Zeta*. Elle est donnée par

$$\mu(T) = \sum_{S \subseteq T} m_\mu(S), \quad \forall T \subseteq N. \quad (1.1)$$

La représentation *duale* d'une fonction d'ensemble est une autre représentation équivalente couramment utilisée. Pour toute fonction d'ensemble μ sur N , la représentation duale de μ , notée $\bar{\mu}$, est définie par

$$\bar{\mu}(A) := \mu(N) - \mu(N \setminus A), \quad \forall A \subseteq N.$$

Il est facile de vérifier que $\bar{\bar{\mu}} = \mu$ si $\mu(\emptyset) = 0$.

Dans le contexte de l'aide multicritère à la décision fondée sur les intégrales non additives et de la théorie des jeux coopératifs, d'autres représentations d'une fonction d'ensemble μ sur N ont été proposées (Grabisch, Marichal & Roubens 2000). Nous donnons ci-après la représentation

en terme d'indices d'interaction de Shapley, définie par

$$I_\mu(S) := \sum_{T \subseteq N \setminus S} \frac{(n-t-s)!t!}{(n-s+1)!} \sum_{L \subseteq S} (-1)^{s-l} \mu(L \cup T), \quad \forall S \subseteq N.$$

Les indices d'interaction ont été introduits afin de mesurer l'interaction entre joueurs dans le cadre d'un jeu coopératif (Grabisch & Roubens 1999) ou encore l'interaction entre critères dans le cadre d'un problème d'aide multicritère à la décision fondée sur les intégrales non additives. Ces quantités seront étudiées plus en détail dans la Section 2.3.

1.2.4 Jeux et capacités

Définition 1.2.2 Une fonction d'ensemble $\mu : \mathcal{P}(N) \rightarrow \mathbb{R}$ est un jeu coopératif sur N à utilité transférable sous forme caractéristique, ou plus simplement un jeu sur N , si $\mu(\emptyset) = 0$.

La définition ci-dessus est à la base de la *théorie des jeux coopératifs* (voir p. ex. Curiel 1997, Peleg & Sudhölter 2003). Dans la suite, l'ensemble des jeux sur N sera noté \mathcal{G}_N . Lorsque l'ensemble N représente un ensemble de votants comme en théorie des votes, $\mu(S) \in \{0, 1\}$ pour tout $S \subseteq N$ et $\mu(S) = 1$ pour une coalition $S \subseteq N$ si et seulement si la motion soumise au vote est acceptée lorsque les votants de S votent pour et $\mu(S) = 0$ sinon. La fonction μ est alors appelée *jeu de vote*. Dans le cas où l'ensemble N représente un ensemble de joueurs, d'agents, d'entreprises, ..., $\mu(S)$ représente le *gain*, généralement financier, que la coalition S pourrait espérer si les acteurs de S décidaient de collaborer. Il est sous-entendu que, finalement, seule la *grande coalition* N sera formée.

Définition 1.2.3 Une fonction d'ensemble $\mu : \mathcal{P}(N) \rightarrow \mathbb{R}$ est une capacité sur N s'il s'agit d'un jeu monotone, c.-à-d. si $\mu(\emptyset) = 0$ et, si, pour tous $S, T \subseteq N$, $S \subseteq T$ implique $\mu(S) \leq \mu(T)$.

La notion de capacité a été introduite par Choquet (1953) dans le contexte de sa *théorie des capacités*. Un concept similaire a été proposé par Sugeno (1974) sous le nom de *mesure floue*. Une capacité est dite normalisée si $\mu(N) = 1$. Dans la suite, l'ensemble des capacités sur N sera noté \mathcal{C}_N et l'ensemble des capacités normalisées \mathcal{C}_N^* .

Dans le contexte de l'aide multicritère à la décision fondée sur les intégrales non additives, $\mu(S)$ peut être vu comme le *poids* ou l'*importance* du sous-ensemble $S \subseteq N$ de critères dans la décision. Plus précisément, $\mu(S)$ s'interprète comme le score global de l'alternative (*binaire*) ayant pour vecteur de scores partiels $(1_S, 0_{N \setminus S})$ (voir p. ex. Labreuche & Grabisch 2003).

Notons qu'il n'existe qu'une capacité normalisée sur N qui soit à la fois additive et cardinale. Nous l'appellerons la *capacité uniforme* et la noterons μ^* . Il est facile de vérifier que μ^* est donnée par

$$\mu^*(T) = t/n, \quad \forall T \subseteq N.$$

La capacité uniforme joue un rôle particulier, notamment dans le contexte de l'aide multicritère à la décision fondée sur les intégrales non additives, car, comme cela pourra être vérifié dans la Section 1.3.1, l'intégrale de Choquet par rapport à μ^* n'est rien d'autre que la moyenne arithmétique simple.

Nous concluons ce paragraphe par un résultat dû à Chateauneuf & Jaffray (1989) relatif aux contraintes de monotonie sur la représentation de Möbius d'une capacité.

Proposition 1.2.1 Une fonction d'ensemble $m : \mathcal{P}(N) \rightarrow \mathbb{R}$ correspond à la représentation de Möbius d'une capacité si et seulement si

$$(i) \quad m(\emptyset) = 0;$$

$$(ii) \quad \sum_{K \subseteq S} m(i \cup K) \geq 0, \text{ pour tout } i \in N, \text{ pour tout } S \subseteq N \setminus i.$$

La fonction m correspond à la représentation de Möbius d'une capacité normalisée si et seulement si, additionnellement, $\sum_{K \subseteq N} m(K) = 1$.

Le résultat ci-dessus est très utile, notamment dans le contexte de l'identification de capacités comme nous le verrons dans le Chapitre 3.

1.2.5 Fonctions d'ensemble k -additives

Définir une fonction d'ensemble μ sur N nécessite de déterminer les 2^n nombres réels $(\mu(S))_{S \subseteq N}$. L'utilisation de fonctions d'ensemble dans les applications se trouve évidemment freinée par cette complexité exponentielle. Dès que n devient important, il est ainsi très courant de considérer que la fonction d'ensemble est additive, au détriment de la qualité de la modélisation. La notion fondamentale de k -additivité, proposée par Grabisch (1997b), permet d'éviter d'avoir à recourir à une simplification si drastique tout en diminuant la complexité de représentation de la fonction d'ensemble.

Afin d'introduire la notion de k -additivité, il est utile de définir celle de *fonction pseudo-booléenne* associée à une fonction d'ensemble.

Définition 1.2.4 Une fonction définie sur $\{0, 1\}^n$ à valeurs dans \mathbb{R} est dite *pseudo-booléenne*.

Toute fonction pseudo-booléenne peut être associée de façon unique à une fonction d'ensemble réelle définie sur $\mathcal{P}(N)$. En effet, à toute fonction $\mu : \mathcal{P}(N) \rightarrow \mathbb{R}$ peut être associée la fonction pseudo-booléenne

$$f^\mu(1_S, 0_{N \setminus S}) := \mu(S), \quad \forall S \subseteq N.$$

et réciproquement, à toute fonction pseudo-booléenne f peut être associée la fonction d'ensemble

$$\mu^f(S) := f(1_S, 0_{N \setminus S}), \quad \forall S \subseteq N.$$

Hammer & Rudeanu (1968) ont démontré que toute fonction pseudo-booléenne f pouvait s'écrire comme un polynôme multilinéaire de n variables de la façon suivante :

$$f(x) = \sum_{T \subseteq N} m_T \prod_{i \in T} x_i, \quad \forall x \in \{0, 1\}^n. \quad (1.2)$$

À partir l'Eq. (1.1), il est facile de vérifier que les coefficients $(m_T)_{T \subseteq N}$ dans l'expression ci-dessus sont les coefficients de la représentation de Möbius m_μ de la fonction d'ensemble μ associée à f .

Si μ est additive, elle est complètement définie par la donnée de la famille $(\mu(i))_{i \in N}$. Il est alors facile de vérifier que sa fonction pseudo-booléenne f^μ est donnée par

$$f^\mu(x) = \sum_{i \in N} m_\mu(i) x_i, \quad \forall x \in \{0, 1\}^n,$$

et qu'elle est donc linéaire. Par extension, Grabisch (1997b) a défini la notion de k -additivité comme suit.

Définition 1.2.5 Une fonction d'ensemble μ définie sur $\mathcal{P}(N)$ est k -additive lorsque sa fonction pseudo-booléenne associée est un polynôme multilinéaire de degré k , c.-à-d. lorsque sa représentation de Möbius satisfait $m_\mu(T) = 0$ pour tout T tel que $|T| > k$ et il existe au moins un sous-ensemble T de cardinal k tel que $m_\mu(T) \neq 0$.

Afin de mieux comprendre l'importance de ce concept, considérons par exemple un jeu μ 2-additif. Pour tout $S \subseteq N$, nous avons alors

$$\mu(S) = \sum_{i \in S} m_\mu(i) + \sum_{ij \subseteq S} m_\mu(ij).$$

Pour les singletons i de N , nous obtenons immédiatement $\mu(i) = m_\mu(i)$ et pour les paires $ij \subseteq N$, nous avons

$$\begin{aligned} \mu(ij) &= m_\mu(i) + m_\mu(j) + m_\mu(ij) \\ &= \mu(i) + \mu(j) + m_\mu(ij). \end{aligned}$$

Ainsi, pour tout $S \subseteq N$, $s \geq 2$, nous obtenons

$$\mu(S) = \sum_{ij \subseteq S} \mu(ij) - (s-2) \sum_{i \in S} \mu(i).$$

Un jeu 2-additif est donc complètement déterminé par ses valeurs sur les singletons et sur les paires de N , soit par $n + \binom{n}{2} = n(n+1)/2$ coefficients. Plus généralement, un jeu k -additif, $1 \leq k \leq n$, est parfaitement défini par la donnée de $\sum_{l=1}^k \binom{n}{l}$ coefficients, par exemple par l'une des trois familles $(\mu(S))_{S \subseteq N, 0 < |S| \leq k}$, $(I_\mu(S))_{S \subseteq N, 0 < |S| \leq k}$ ou $(m_\mu(S))_{S \subseteq N, 0 < |S| \leq k}$. Dans ce dernier cas, nous avons immédiatement

$$\mu(S) = \sum_{\substack{T \subseteq S \\ 0 < |T| \leq k}} m_\mu(T), \quad \forall S \subseteq N. \quad (1.3)$$

La notion de k -additivité est d'une importance fondamentale en pratique, notamment dans le contexte de l'identification de modèles décisionnels fondés sur l'intégrale de Choquet (cf. Chapitre 3).

1.2.6 Le treillis $(\mathcal{P}(N), \subseteq)$

Du point de vue de la théorie des ensembles ordonnés, une capacité est simplement une fonction isotone du treillis $(\mathcal{P}(N), \subseteq)$ dans (\mathbb{R}^+, \leq) . À titre d'exemple, le diagramme de Hasse de $(\mathcal{P}(N), \subseteq)$ est donné dans la Figure 1.1 pour $n = 3$. Il est alors commode de représenter la capacité relativement aux chaînes maximales du treillis $(\mathcal{P}(N), \subseteq)$. Comme nous le verrons dans la Section 1.3.1, dans le contexte de l'agrégation par intégrale de Choquet, cette représentation est très utile même dans le cas de jeux ou de fonctions d'ensemble.

Soit $\mathcal{M}_{(\mathcal{P}(N), \subseteq)}$ l'ensemble des chaînes maximales du treillis $(\mathcal{P}(N), \subseteq)$ et soit Π_N l'ensemble des permutations sur N . Les ensembles $\mathcal{M}_{(\mathcal{P}(N), \subseteq)}$ et Π_N sont clairement équipollents. En effet, à chaque permutation $\sigma \in \Pi_N$ peut être associée une unique chaîne maximale $c_\sigma \in \mathcal{M}_{(\mathcal{P}(N), \subseteq)}$ définie par

$$c_\sigma := (\emptyset \subsetneq \{\sigma(n)\} \subsetneq \{\sigma(n-1), \sigma(n)\} \subsetneq \cdots \subsetneq \{\sigma(1), \dots, \sigma(n)\} = N).$$

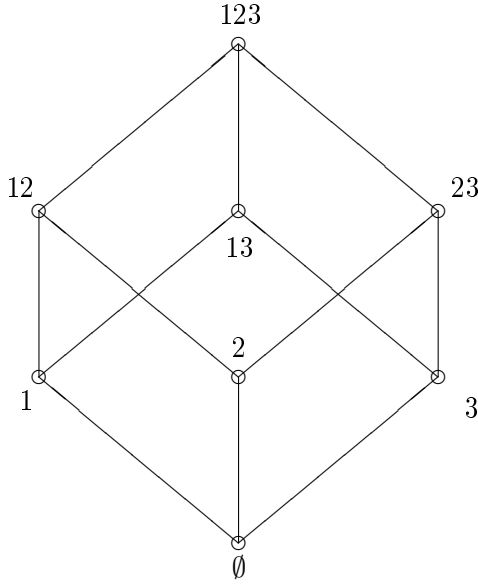


FIG. 1.1 – Diagramme de Hasse de $(\mathcal{P}(N), \subseteq)$ pour $n = 3$.

Étant donnée une fonction d'ensemble μ définie sur $\mathcal{P}(N)$, à chaque permutation $\sigma \in \Pi_N$, nous associons la distribution de réels ω_σ^μ sur N définie par

$$\omega_\sigma^\mu(i) := \mu(\{\sigma(i), \dots, \sigma(n)\}) - \mu(\{\sigma(i+1), \dots, \sigma(n)\}), \quad \forall i \in N. \quad (1.4)$$

De façon équivalente, à une chaîne maximale $c_\sigma \in \mathcal{M}(\mathcal{P}(N), \subseteq)$, nous associons la distribution $\omega_{c_\sigma}^\mu := \omega_\sigma^\mu$.

Nous dirons dans la suite que la distribution ω_σ^μ est *induite* par la chaîne c_σ de $(\mathcal{P}(N), \subseteq)$.

Une fonction d'ensemble μ sur N peut ainsi être vue comme un ensemble de $n!$ distributions correspondant aux $n!$ chaînes maximales de $(\mathcal{P}(N), \subseteq)$. Notons que si μ est une capacité, alors, par monotonie, les nombres $\omega_\sigma^\mu(i)$ sont nécessairement non négatifs pour tout $\sigma \in \Pi_N$ et tout $i \in N$. Si, de plus, μ est normalisée, alors ω_σ^μ est une distribution de probabilité sur N pour tout $\sigma \in \Pi_N$.

1.3 L'intégrale de Choquet discrète en tant qu'opérateur d'agrégation

Le concept d'*intégrale par rapport à une mesure* (voir p. ex. Rudin 1975) a été étendu aux mesures non additives (voir p. ex. Choquet 1953, Sugeno 1974, Denneberg 1994, Denneberg 2000, Murofushi & Sugeno 2000, Benvenuti, Mesiar & Vivona 2002). Dans un contexte discret, ces intégrales généralisées, dont l'intégrale de Choquet et l'intégrale de Sugeno sont les représentantes les plus connues, peuvent être vues comme des opérateurs d'agrégation particulièrement flexibles, ce qui suggère de les utiliser pour la représentation numérique des préférences en aide multicritère à la décision (Grabisch 1992, Marichal 2000a, Grabisch & Labreuche 2004). L'intégrale de Choquet est par ailleurs à la base de nombreux modèles de décision dans l'incertain et le risque (voir p. ex. Schmeidler 1989, Tversky & Kahneman 1992, Chateauneuf 1994, Narukawa & Murofushi 2004).

Nous nous limitons dans ce paragraphe à une présentation de l'intégrale de Choquet dans le cas discret en mettant l'accent sur son rôle en agrégation.

1.3.1 Définition

Dans la suite, toute fonction $x : N \rightarrow \mathbb{R}$ sera assimilée à un vecteur de \mathbb{R}^n avec la notation $x_i := x(i)$ pour tout $i \in N$.

Définition 1.3.1 Soit μ un jeu sur N et soit x une fonction de N dans \mathbb{R}^+ . L'intégrale de Choquet de x par rapport à μ est définie par

$$C_\mu(x) := \sum_{i=1}^n \omega_\sigma^\mu(i) x_{\sigma(i)},$$

où σ est une permutation sur N telle que $x_{\sigma(1)} \leq \dots \leq x_{\sigma(n)}$.

Il est facile de vérifier que l'intégrale de Choquet par rapport à un jeu est une fonction linéaire sur chaque cône

$$\mathcal{K}_\sigma := \{x \in (\mathbb{R}^+)^n \mid x_{\sigma(1)} \leq \dots \leq x_{\sigma(n)}\} \quad (\sigma \in \Pi_N).$$

En nous reportant à la Section 1.2.6, nous pouvons voir que, pour $x \in \mathcal{K}_\sigma$, les coefficients des évaluations partielles $x_{\sigma(1)}, \dots, x_{\sigma(n)}$ sont induits par la chaîne maximale c_σ de $(\mathcal{P}(N), \subseteq)$.

Le résultat suivant, dû à Chateauneuf & Jaffray (1989), donne l'expression de l'intégrale de Choquet en fonction de la représentation de Möbius d'un jeu.

Proposition 1.3.1 Soit μ un jeu sur N et soit x une fonction de N dans \mathbb{R}^+ . En fonction de la représentation de Möbius m_μ de μ , l'intégrale de Choquet de x par rapport à μ est donnée par

$$C_\mu(x) = \sum_{T \subseteq N} m_\mu(T) \bigwedge_{i \in T} x_i, \quad (1.5)$$

où $\bigwedge_{i \in T} x_i$ désigne le minimum des x_i , $i \in T$.

Il apparaît ainsi clairement que l'intégrale de Choquet par rapport à un jeu est une fonction continue sur $(\mathbb{R}^+)^n$.

1.3.2 Caractérisation axiomatique en tant qu'opérateur d'agrégation

Dans le contexte de l'agrégation, l'intégrale de Choquet est généralement définie par rapport à une capacité normalisée. Le cadre est alors souvent celui de l'aide multicritère à la décision, en particulier celui de la théorie de l'utilité multiattribut (voir p. ex. Keeney & Raiffa 1976). Dans ce cas, N désigne un ensemble de critères servant à décrire un ensemble d'alternatives $\mathcal{A} := \{a, b, c, \dots\}$ parmi lesquelles, typiquement, un décideur doit choisir. Chaque alternative $a \in \mathcal{A}$ est identifiée à son vecteur de *scores partiels* ou *utilités* $(a_1, \dots, a_n) \in \mathbb{R}^n$. Les scores partiels sont de plus ici supposés définis sur la même échelle d'intervalle, typiquement $[0, 1]$. Le rôle de l'opérateur d'agrégation, fonction de \mathbb{R}^n dans \mathbb{R} , est alors d'associer à chaque alternative un score global calculé à partir de ses scores partiels tout en prenant en compte les *préférences*

du décideur. Le cadre précis de la théorie de l'utilité multiattribut sera décrit de façon plus rigoureuse dans le Chapitre 3.

En tant qu'opérateur d'agrégation, l'intégrale de Choquet par rapport à une capacité normalisée satisfait des propriétés particulièrement remarquables (voir p. ex. Grabisch 1997a, Marichal 2000a, Labreuche & Grabisch 2003). Elle est par exemple continue, non décroissante par rapport à chacun de ses arguments², comprise entre le minimum et le maximum des scores partiels, stable pour les transformations affines positives (transformations admissibles) de l'échelle des scores partiels et coïncide avec la moyenne arithmétique pondérée dès que la capacité est additive.

Une caractérisation axiomatique de l'intégrale de Choquet en tant qu'opérateur d'agrégation a été proposée par Marichal (2000a).

Théorème 1.3.1 *Les opérateurs $M_\mu : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mu \in \mathcal{C}_N^*$, sont :*

- (i) *linéaires par rapport à la capacité ;*
- (ii) *non décroissants par rapport à chaque argument ;*
- (iii) *stables pour les transformations linéaires admissibles, c.-à-d.*

$$M_\mu(rx + s1_N) = rM_\mu(x) + s$$

pour tous $x \in \mathbb{R}^n$, $r > 0$, $s \in \mathbb{R}$ et pour tout $\mu \in \mathcal{C}_N^$;*

- (iv) *proprement pondérés par μ , c.-à-d.*

$$M_\mu(1_S, 0_{N \setminus S}) = \mu(S), \quad \forall S \subseteq N, \forall \mu \in \mathcal{C}_N^* ;$$

si et seulement si $M_\mu = C_\mu$ pour tout $\mu \in \mathcal{C}_N^$.*

Les axiomes précédents sont très naturels dans le contexte de l'agrégation. Le premier axiome impose que la fonction d'agrégation soit la plus simple possible. Le second axiome stipule qu'une augmentation d'un score partiel sur un critère ne doit pas entraîner une diminution du score global. Le troisième axiome requiert que l'agrégation soit "stable" dans le cas de n'importe quel changement admissible d'échelle. Enfin, si l'échelle d'évaluation est $[0, 1]$, le dernier axiome impose que le poids du sous-ensemble S de critères soit pris égal à l'évaluation globale de l'alternative dont le vecteur de scores partiels est $(1_S, 0_{N \setminus S})$, c.-à-d. l'alternative, dite *binaire*, parfaitement satisfaisante selon les critères de S et totalement insatisfaisante selon les critères restants.

Dans le contexte de l'agrégation, l'intérêt pratique de l'intégrale de Choquet par rapport à une capacité normalisée est qu'elle généralise un nombre important d'opérateurs classiques (voir p. ex. Marichal 2000b). L'intégrale de Choquet par rapport à une capacité normalisée additive n'est rien d'autre qu'une moyenne arithmétique pondérée. Lorsque la capacité normalisée est cardinale, l'intégrale de Choquet est simplement une combinaison linéaire de statistiques d'ordre, appelée *ordered weighted averaging operators (OWA)* par Yager (1988), et généralisant entre autres le minimum et le maximum. Enfin, par rapport à une capacité *binaire*, c.-à-d. à valeurs dans $\{0, 1\}$, l'intégrale de Choquet n'est rien d'autre qu'un *polynôme latticiel* (Marichal 2005, Marichal & Kojadinovic 2006), dont le minimum et le maximum partiels sont des cas particuliers.

1.3.3 Interprétation géométrique

Une interprétation géométrique remarquable de l'intégrale de Choquet a été récemment donnée par Grabisch (2004) (voir aussi Grabisch 2005c, Grabisch 2005a). Soit M une fonction de

²Cette propriété est une conséquence directe de la monotonie de la capacité.

$[0, 1]^n$ dans \mathbb{R} et soit μ un jeu sur N tel que $M(1_S, 0_{N \setminus S}) = \mu(S)$ pour tout $S \subseteq N$. La fonction M est donc connue sur les sommets de l'hypercube unité. Grabisch (2004) s'est intéressé à la détermination de M sur $[0, 1]^n \setminus \{0, 1\}^n$ par interpolation linéaire. Il a montré que l'intégrale de Choquet correspond à l'unique interpolation linéaire utilisant $n + 1$ sommets de $\{0, 1\}^n$. En ce sens, il s'agit de l'interpolation linéaire la plus simple.

Dans le contexte de l'aide multicritère à la décision, l'intégrale de Choquet peut ainsi être vue comme le moyen le plus simple pour étendre à des alternatives quelconques le raisonnement d'un décideur sur des alternatives *binaires*.

1.3.4 Lien avec l'extension de Lovász

Il est intéressant de remarquer que l'intégrale de Choquet par rapport à un jeu μ coïncide sur $[0, 1]^n$ avec l'*extension* de Lovász (1983) de sa fonction pseudo-booléenne associée f^μ (voir p. ex. Grabisch 2000, Marichal 2002a).

Une extension d'une fonction pseudo-booléenne $f : \{0, 1\}^n \rightarrow \mathbb{R}$ est une fonction g définie sur l'hypercube $[0, 1]^n$ qui coïncide avec f sur $\{0, 1\}^n$. L'extension de Lovász de f , notée \hat{f} , est définie par

$$\hat{f}(x) := \sum_{T \subseteq N} m_T \bigwedge_{i \in T} x_i, \quad \forall x \in [0, 1]^n,$$

où les coefficients m_T sont les coefficients de la forme multilinéaire de f donnée par l'Eq. (1.2).

Ce lien direct entre l'intégrale de Choquet et l'extension de Lovász suggère l'interprétation suivante : l'intégrale de Choquet par rapport à un jeu μ peut être considérée comme une extension de μ à l'ensemble de sous-ensembles flous de N (Grabisch 2000).

1.3.5 Premières extensions de l'intégrale de Choquet

Comme nous pouvons le voir en considérant la Définition 1.3.1, l'intégrale de Choquet est définie pour les fonctions non négatives uniquement. Pour les fonctions pouvant prendre des valeurs négatives, il existe plusieurs extensions de l'intégrale de Choquet (voir p. ex. Denneberg 1994, Grabisch & Labreuche 2002). Pour tout $x \in \mathbb{R}^n$, x^+ (resp. x^-) désigne la partie positive (resp. négative) de x , définie par $x^+ := x \vee 0$ (resp. $x^- := (-x) \vee 0$), où \vee désigne le maximum.

Une première généralisation de l'intégrale de Choquet est connue sous le nom d'*intégrale de Choquet symétrique*, aussi appelée *intégrale de Šipoš* (1979). Elle est définie, pour tout jeu μ sur N , par

$$\check{S}_\mu(x) := C_\mu(x^+) - C_\mu(x^-), \quad \forall x \in \mathbb{R}^n.$$

Une seconde généralisation est connue sous le nom d'*intégrale de Choquet asymétrique*. Elle est définie, pour tout jeu μ sur N , par

$$\bar{C}_\mu(x) := C_\mu(x^+) - C_{\bar{\mu}}(x^-), \quad \forall x \in \mathbb{R}^n.$$

Cette dernière définition coïncide avec la définition usuelle (cf. Définition 1.3.1) étendue aux intégrandes réelles dans le sens où $\bar{C}_\mu(x) = C_\mu(x)$ pour tout $x \in \mathbb{R}^n$.

Dans le contexte de l'agrégation, l'intégrale de Choquet symétrique peut être vue comme une généralisation permettant d'agréger des évaluations définies sur une échelle *bipolaire* (Grabisch & Labreuche 2002).

Comparée à une échelle classique (*unipolaire*), une échelle bipolaire est caractérisée par l'existence d'un *niveau neutre* tel que les évaluations supérieures à ce niveau sont ressenties comme “bonnes” ou “positives” par le décideur, alors que les évaluations inférieures au niveau neutre sont considérées comme “mauvaises” ou “négatives” (voir p. ex. Grabisch 2005b, Section 3.3 pour une présentation plus approfondie du concept d'*échelle bipolaire*). En plus du niveau neutre, l'existence de deux autres niveaux est supposée : un niveau *totalelement satisfaisant* et un niveau *totalelement inacceptable*. Le prototype de l'échelle bipolaire est l'intervalle $[-1, 1]$, la valeur 0 jouant le rôle du niveau neutre, la valeur 1 le rôle du niveau totalement satisfaisant et la valeur -1 celui du niveau totalement inacceptable.

Utiliser en revanche l'intégrale de Choquet (asymétrique) dans un contexte bipolaire revient à ignorer le niveau neutre de l'échelle. En fait, aussi bien l'intégrale de Choquet symétrique que l'intégrale de Choquet asymétrique sont des cas particuliers du modèle CPT (*Cumulative Prospect Theory*) introduit par Tversky & Kahneman (1992), lequel, comme nous le verrons dans la section suivante, est un cas particulier de l'intégrale de Choquet par rapport à un jeu bi-coopératif (Grabisch & Labreuche 2005b, Grabisch & Labreuche 2005c).

1.4 Généralisations

Les notions de *jeu coopératif* et de *capacité* définies dans la section précédente ont été récemment généralisées afin de permettre une modélisation plus adéquate de la réalité. Une synthèse des nombreuses extensions proposées dans la littérature peut être trouvée dans (Grabisch 2005a). Nous nous limitons dans cette section à une présentation des notions de *jeu bi-coopératif* et de *bi-capacité*.

1.4.1 Jeux bi-coopératifs et bi-capacités

Soit $\mathcal{Q}(N) := \{(A, B) \in \mathcal{P}(N) \times \mathcal{P}(N) \mid A \cap B = \emptyset\}$.

Définition 1.4.1 Une fonction $v : \mathcal{Q}(N) \rightarrow \mathbb{R}$ est un jeu bi-coopératif sur N si $v(\emptyset, \emptyset) = 0$.

La notion de jeu bi-coopératif a été définie par Bilbao (2000) et permet de modéliser les situations dans lesquelles un joueur $i \in N$ peut *collaborer positivement*, *collaborer négativement* ou *ne pas participer* dans le cadre d'un jeu. Une *bi-coalition* $(A, B) \in \mathcal{Q}(N)$ représente ainsi la situation où les joueurs de A collaborent positivement, ceux de B négativement, les joueurs restants s'abstenant de participer. La quantité $v(A, B)$ quantifie alors typiquement le *gain* des joueurs de A dans la situation (A, B) (Labreuche & Grabisch 2006b).

Les jeux bi-coopératifs peuvent être vus comme des extensions des *jeux de vote ternaires* introduits par Felsenthal & Machover (1997). Dans ce contexte, un votant $i \in N$ peut voter pour une motion, contre cette motion ou s'abstenir. Une *bi-coalition* $(A, B) \in \mathcal{Q}(N)$ représente la situation où les votants de A votent pour la motion, ceux de B contre, le reste des votants s'abstenant. Un jeu de vote ternaire v est une fonction de $\mathcal{Q}(N)$ dans $\{-1, 1\}$ telle que $v(A, B) = 1$ (resp. $v(A, B) = -1$) si et seulement si la motion est acceptée (resp. rejetée) dans la situation (A, B) .

L'ensemble des jeux bi-coopératifs sur N sera noté \mathcal{BG}_N dans la suite.

Définition 1.4.2 Une fonction $v : \mathcal{Q}(N) \rightarrow \mathbb{R}$ est une bi-capacité sur N s'il s'agit d'un jeu bi-coopératif et si, pour tous $A, B \subseteq N$, $A \subseteq B$ implique $v(A, \cdot) \leq v(B, \cdot)$ et $v(\cdot, A) \geq v(\cdot, B)$.

Une bi-capacité v est dite normalisée si, additionnellement, $v(N, \emptyset) = 1$ et $v(\emptyset, N) = -1$.

Les bi-capacités ont été introduites par Grabisch & Labreuche (2005a) dans le cadre de l'aide multicritère à la décision fondée sur les intégrales non additives afin de permettre la prise en compte d'évaluations partielles données sur une échelle *bipolaire*. Dans ce contexte, une bi-coalition $(A, B) \in \mathcal{Q}(N)$ correspond à la situation où les critères de A sont satisfaits positivement et ceux de B négativement. En considérant $[-1, 1]$ comme prototype d'échelle bipolaire, la valeur $v(A, B)$ s'interprète comme l'évaluation de l'alternative, dite *ternaire*, dont les évaluations partielles sont représentées par le vecteur $(1_A, -1_B, 0_{N \setminus (A \cup B)}) := (1_A, 0_{N \setminus A}) - (1_B, 0_{N \setminus B})$ de $[-1, 1]^n$.

Noter qu'une autre vision de la bipolarité, dite *double unipolaire* (voir p. ex. Grabisch 2005b, Section 3.1) a conduit Greco, Matarazzo & Slowinski (2002) à proposer la notion de *capacité bipolaire* (voir aussi Greco, Matarazzo & Slowinski 2003).

Dans la suite, l'ensemble des bi-capacités sur N sera noté \mathcal{BC}_N et l'ensemble des bi-capacités normalisées \mathcal{BC}_N^* .

Un jeu bi-coopératif v sur N est dit :

- de type CPT (*Cumulative Prospect Theory*) s'il existe deux jeux μ_1, μ_2 sur N tels que

$$v(A, B) = \mu_1(A) - \mu_2(B), \quad \forall (A, B) \in \mathcal{Q}(N).$$

Lorsque $\mu_2 = \mu_1$ (resp. $\mu_2 = \bar{\mu}_1$), le jeu bi-coopératif est en plus dit *symétrique* (resp. *asymétrique*).

- *additif* s'il est de type CPT avec μ_1, μ_2 additives, c.-à-d.

$$v(A, B) = \sum_{i \in A} \mu_1(i) - \sum_{i \in B} \mu_2(i), \quad \forall (A, B) \in \mathcal{Q}(N).$$

Noter qu'un jeu bi-coopératif additif avec $\mu_1 = \mu_2$ est à la fois symétrique et asymétrique car $\bar{\mu}_1 = \mu_1$.

Dans la suite, pour indiquer qu'un jeu bi-coopératif de type CPT est défini à partir de deux jeux μ_1, μ_2 , nous le noterons v_{μ_1, μ_2} .

Un jeu bi-coopératif v sur N sera dit *différence cardinal* si, pour tout $(A, B) \in \mathcal{Q}(N)$, $v(A, B)$ ne dépend que de la différence $a - b$ de cardinaux. Il est facile de vérifier qu'il n'existe qu'une bi-capacité normalisée sur N qui soit à la fois additive et différence cardinale. Nous l'appellerons la bi-capacité *uniforme* et la noterons v^* . Elle est définie par

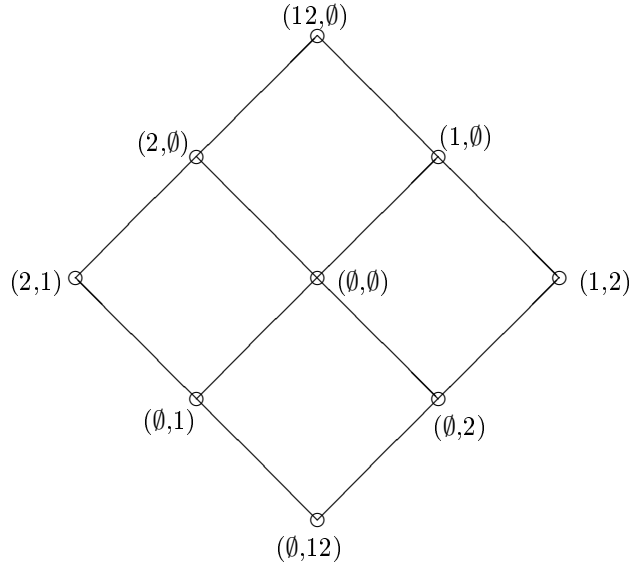
$$v^*(A, B) := \frac{a - b}{n} = \mu^*(A) - \mu^*(B), \quad \forall (A, B) \in \mathcal{Q}(N).$$

La bi-capacité uniforme joue un rôle similaire à celui de la capacité uniforme μ^* définie dans la Section 1.2.4 car, comme cela pourra être vérifié dans la Section 1.4.4, l'intégrale de Choquet par rapport à v^* n'est rien d'autre que la moyenne arithmétique simple.

1.4.2 L'inf-demi-trellis $(\mathcal{Q}(N), \sqsubseteq)$

Dans leurs travaux fondateurs sur les bi-capacités, Grabisch & Labreuche (2005a) ont défini l'ordre suivant sur $\mathcal{Q}(N)$:

$$(A, B) \sqsubseteq (A', B') \iff A \subseteq A' \text{ et } B \supseteq B'.$$

FIG. 1.2 – Diagramme de Hasse de $(\mathcal{Q}(N), \subseteq)$ pour $n = 2$.

L'ensemble ordonné $(\mathcal{Q}(N), \subseteq)$ est alors un treillis dont le plus petit élément est (\emptyset, N) et le plus grand (N, \emptyset) , l'élément (\emptyset, \emptyset) ne jouant pas de rôle particulier. Le diagramme de Hasse de $(\mathcal{Q}(N), \subseteq)$ est donné dans la Figure 1.2 pour $n = 2$.

Les derniers développements dans le domaine (voir p. ex. Grabisch & Labreuche 2005c, Grabisch & Labreuche 2005d, Fujimoto & Murofushi 2005, Kojadinovic & Marichal 2006, Kojadinovic 2006c, Labreuche & Grabisch 2006b) suggèrent que l'ordre le plus approprié est simplement l'ordre produit :

$$(A, B) \subseteq (A', B') \iff A \subseteq A' \text{ et } B \subseteq B'.$$

Comme nous pouvons le voir sur la Figure 1.3, l'ensemble ordonné $(\mathcal{Q}(N), \subseteq)$ est alors un inf-demi-treillis dont le plus petit élément est (\emptyset, \emptyset) (voir Grabisch & Labreuche 2005c, pour plus de détails).

Soit $\mathcal{M}_{(\mathcal{Q}(N), \subseteq)}$ l'ensemble des chaînes maximales de $(\mathcal{Q}(N), \subseteq)$. Il est facile de vérifier que les ensembles $\mathcal{M}_{(\mathcal{Q}(N), \subseteq)}$ et $\mathcal{P}(N) \times \Pi_N$ sont équipollents, ce qui implique que $|\mathcal{M}_{(\mathcal{Q}(N), \subseteq)}| = 2^n n!$. En effet, à chaque couple composé d'un sous-ensemble $N^+ \subseteq N$ et d'une permutation $\sigma \in \Pi_N$, peut être associée une unique chaîne maximale $c_{N^+, \sigma} \in \mathcal{M}_{(\mathcal{Q}(N), \subseteq)}$ définie par

$$\begin{aligned} c_{N^+, \sigma} := & \left((\emptyset, \emptyset) \subsetneq (\{\sigma(n)\} \cap N^+, \{\sigma(n)\} \cap N^-) \right. \\ & \subsetneq (\{\sigma(n-1), \sigma(n)\} \cap N^+, \{\sigma(n-1), \sigma(n)\} \cap N^-) \\ & \subsetneq \dots \\ & \left. \subsetneq (\{\sigma(1), \dots, \sigma(n)\} \cap N^+, \{\sigma(1), \dots, \sigma(n)\} \cap N^-) = (N^+, N^-) \right), \end{aligned}$$

avec $N^- := N \setminus N^+$.

En procédant comme dans la Section 1.2.6, étant donné un jeu bi-coopératif v sur N , à chaque couple composé d'un sous-ensemble $N^+ \subseteq N$ et d'une permutation $\sigma \in \Pi_N$ (c.-à-d. à chaque

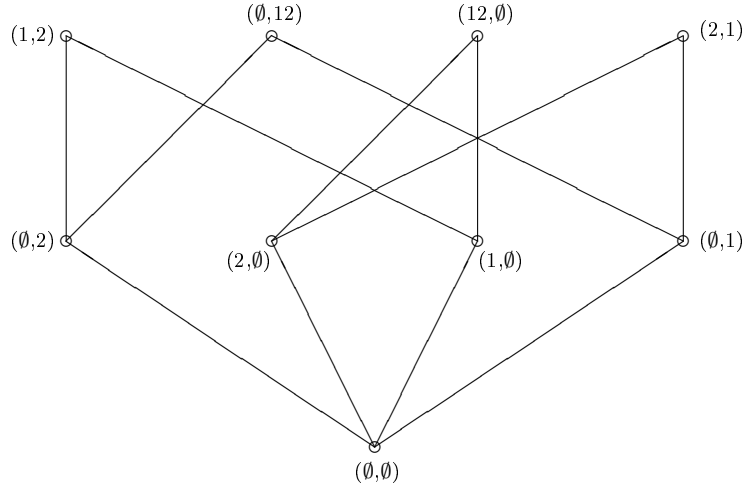


FIG. 1.3 – Diagramme de Hasse de $(\mathcal{Q}(N), \subseteq)$ pour $n = 2$.

chaîne maximale $c_{N+, \sigma}$, nous associons la distribution de réels $\omega_{N+, \sigma}^v$ sur N définie par

$$\omega_{N+, \sigma}^v(i) := \mu_{N+}^v(\{\sigma(i), \dots, \sigma(n)\}) - \mu_{N+}^v(\{\sigma(i+1), \dots, \sigma(n)\}), \quad \forall i \in N,$$

où μ_{N+}^v est un jeu sur N défini par

$$\mu_{N+}^v(C) := v(C \cap N^+, C \cap N^-), \quad \forall C \subseteq N. \quad (1.6)$$

De façon équivalente,

$$\omega_{N+, \sigma}^v(i) = v(c_{N+, \sigma}[n - i + 1]) - v(c_{N+, \sigma}[n - i]), \quad \forall i \in N,$$

où, pour tout $i \in N$, $c_{N+, \sigma}[i]$ désigne le i -ème élément de la chaîne maximale $c_{N+, \sigma}$.

Nous dirons dans la suite que la distribution $\omega_{N+, \sigma}^v$ est *induite* par la chaîne maximale $c_{N+, \sigma}$ de $(\mathcal{Q}(N), \subseteq)$.

Il est important de noter qu'en général, $\omega_{N+, \sigma}^v$ n'est ni une distribution de probabilité, ni une distribution de nombres non négatifs. Une exception importante est donnée par le résultat suivant (Kojadinovic & Marichal 2006, Lemme 9) : lorsque v est une bi-capacité asymétrique normalisée, $\omega_{N+, \sigma}^v$ est une distribution de probabilité.

Un jeu bi-coopératif peut ainsi être vu comme un ensemble de $2^n n!$ distributions correspondant aux $2^n n!$ chaînes maximales de $(\mathcal{Q}(N), \subseteq)$. Cette vision d'un jeu bi-coopératif est particulièrement utile dans le contexte de l'agrégation par intégrale de Choquet *bipolaire* comme nous le verrons dans la Section 1.4.4.

1.4.3 Transformée de Möbius sur $(\mathcal{Q}(N), \subseteq)$

En partant de la définition générale de Rota (1964), Grabisch & Labreuche (2005c) (voir aussi Fujimoto & Murofushi 2005) ont déterminé la transformée de Möbius d'un jeu bi-coopératif relativement à l'inf-demi-trellis $(\mathcal{Q}(N), \subseteq)$. Pour un jeu bi-coopératif v sur N , elle est donnée

par

$$m_v(S, T) := \sum_{\substack{S' \subseteq S \\ T' \subseteq T}} (-1)^{|S \setminus S'| + |T \setminus T'|} v(S', T'), \quad \forall (S, T) \in \mathcal{Q}(N).$$

Le jeu bi-coopératif peut être retrouvé à partir de la représentation de Möbius en utilisant la transformation inverse suivante :

$$v(S, T) = \sum_{\substack{S' \subseteq S \\ T' \subseteq T}} m_v(S', T'), \quad \forall (S, T) \in \mathcal{Q}(N).$$

La notion de *jeu bi-coopératif k -additif* peut être définie de façon immédiate par analogie avec la définition “unipolaire” donnée dans la Section 1.2.5 (Fujimoto & Murofushi 2005, Grabisch & Labreuche 2005c).

1.4.4 L'intégrale de Choquet par rapport à un jeu bi-coopératif

Dans le contexte de l'aide multicritère à la décision fondée sur les intégrales non additives et lorsque l'échelle d'évaluation des critères est bipolaire, Grabisch & Labreuche (2005b) ont proposé une extension naturelle de l'intégrale de Choquet qui généralise les intégrales de Choquet symétrique et asymétrique ainsi que le modèle CPT de Tversky & Kahneman (1992).

Définitions et principales propriétés

Définition 1.4.3 *L'intégrale de Choquet d'une fonction $x : N \rightarrow \mathbb{R}$ par rapport à un jeu bi-coopératif v sur N est définie par*

$$C_v(x) := C_{\mu_{N^+}^v}(|x|),$$

où $\mu_{N^+}^v$ est le jeu sur N défini par l'Eq. (1.6) et où $N^+ := \{i \in N \mid x_i \geq 0\}$, $N^- := N \setminus N^+$.

Comme dans le cas unipolaire, l'intégrale de Choquet par rapport à un jeu bi-coopératif est une fonction continue, linéaire sur chaque cône

$$\mathcal{K}_{N^+, \sigma} := \{x \in \mathbb{R}^n \mid x_i \geq 0, \forall i \in N^+, x_i < 0, \forall i \in N^-, \text{ et } |x_{\sigma(1)}| \leq \dots \leq |x_{\sigma(n)}|\}, \quad (N^+ \subseteq N, \sigma \in \Pi_N),$$

l'union de ces cônes couvrant \mathbb{R}^n . En effet, pour chaque $x \in \mathbb{R}^n$, il peut être vérifié qu'il existe $N^+ \subseteq N$ et $\sigma \in \Pi_N$ tels que $x \in \mathcal{K}_{N^+, \sigma}$ et,

$$C_v(x) = \sum_{i \in N} \omega_{N^+, \sigma}^v(i) |x_{\sigma(i)}| = \sum_{\sigma(i) \in N^+} \omega_{N^+, \sigma}^v(i) x_{\sigma(i)} - \sum_{\sigma(i) \in N^-} \omega_{N^+, \sigma}^v(i) x_{\sigma(i)}.$$

En nous reportant à la Section 1.4.2, nous pouvons voir que, pour $x \in \mathcal{K}_{N^+, \sigma}$, les coefficients des évaluations partielles $|x_{\sigma(1)}|, \dots, |x_{\sigma(n)}|$ sont induits par la chaîne maximale $c_{N^+, \sigma}$ de $(\mathcal{Q}(N), \subseteq)$.

La proposition suivante (Fujimoto & Murofushi 2005, Grabisch & Labreuche 2005c) donne l'expression de l'intégrale de Choquet en fonction de la transformée de Möbius sur $(\mathcal{Q}(N), \subseteq)$.

Proposition 1.4.1 *Pour tout jeu bi-coopératif v sur N , nous avons*

$$C_v(x) = \sum_{(S,T) \in \mathcal{Q}(N)} m_v(S,T) \left[\bigwedge_{i \in S} x_i^+ \wedge \bigwedge_{i \in T} x_i^- \right], \quad \forall x \in \mathbb{R}^n.$$

L'intégrale de Choquet par rapport à un jeu bi-coopératif généralise bien les intégrales de Choquet symétrique et asymétrique. En effet, Grabisch & Labreuche (2005b, Prop. 1) ont montré que, pour tout jeu bi-coopératif symétrique $v_{\mu,\mu}$ sur N ,

$$C_{v_{\mu,\mu}}(x) = \check{S}_{\mu}(x), \quad \forall x \in \mathbb{R}^n,$$

et que, pour tout jeu bi-coopératif asymétrique $v_{\mu,\bar{\mu}}$ sur N ,

$$C_{v_{\mu,\bar{\mu}}}(x) = C_{\mu}(x), \quad \forall x \in \mathbb{R}^n. \quad (1.7)$$

Le modèle CPT de Tversky & Kahneman (1992) est quant à lui obtenu en considérant des jeux bi-coopératifs de type CPT.

Une des propriétés les plus importantes de l'intégrale de Choquet par rapport à un jeu bi-coopératif v sur N est que $C_v(1_A, -1_B, 0_{N \setminus (A \cup B)}) = v(A, B)$ pour tout $(A, B) \in \mathcal{Q}(N)$. Cela nous amène à l'interprétation géométrique donnée dans (Grabisch 2004, Grabisch 2005a, Grabisch 2005c) : l'intégrale de Choquet par rapport à un jeu bi-coopératif restreinte à $[-1, 1]^n$ peut être interprétée comme l'interpolation linéaire la plus simple sur $[-1, 1]^n$ d'une fonction connue seulement sur $\{-1, 0, 1\}^n$ (cf. Section 1.3.3).

L'intégrale de Choquet bipolaire en tant qu'opérateur d'agrégation

En tant qu'opérateur d'agrégation, l'intégrale de Choquet bipolaire est utilisée par rapport à des bi-capacités (normalisées), ce qui assure sa non décroissance. Une caractérisation axiomatique dans l'esprit de celle de Marichal (2000a) pour l'intégrale de Choquet par rapport à une capacité normalisée a été proposée par Labreuche & Grabisch (2006a).

Étant donnée une bi-capacité normalisée v sur N et $x \in \mathcal{K}_{N^+, \sigma}$, il peut être vérifié que $\omega_{N^+, \sigma}^v(i)$ et $x_{\sigma(i)}$ sont de même signe pour tout $i \in N$. Il s'ensuit que l'intégrale de Choquet par rapport à v de x peut se récrire sous la forme de la somme pondérée suivante :

$$C_v(x) = \sum_{i \in N} |\omega_{N^+, \sigma}^v(i)| x_{\sigma(i)}.$$

De façon équivalente,

$$C_v(x) = \left(\sum_{i \in N} p_{N^+, \sigma}^v(i) x_{\sigma(i)} \right) \times \left(\sum_{i \in N} |\omega_{N^+, \sigma}^v(i)| \right), \quad (1.8)$$

où

$$p_{\sigma, N^+}^v(i) := \frac{|\omega_{\sigma, N^+}^v(i)|}{\sum_{j \in N} |\omega_{\sigma, N^+}^v(j)|}, \quad \forall i \in N. \quad (1.9)$$

L'expression donnée dans l'Eq. (1.8) est à la base de la *mesure d'uniformité* et des *indices d'importance relative* proposés dans (Kojadinovic & Marichal 2006) et dans (Kojadinovic 2006c) respectivement. Ces notions, permettant d'interpréter l'agrégation par intégrale de Choquet bipolaire, seront présentées dans le chapitre suivant.

Interprétation de mesures non additives



Jacques Rouxel, *Les Shadoks*.

Sommaire

2.1	Introduction	24
2.2	Notion de valeur	25
2.2.1	Valeur de Shapley	25
2.2.2	Valeurs généralisées	27
2.2.3	Indices d'importance dans le cas bipolaire	27
2.2.4	Valeur de Shapley par rapport à un jeu bi-coopératif	31
2.3	Indices d'interaction	31
2.3.1	Notion d'interaction	31
2.3.2	Indices d'interaction de Shapley par rapport à un jeu coopératif	32
2.3.3	Extension au cas bipolaire	33
2.4	Mesures d'uniformité	34
2.4.1	Notion d'entropie	34
2.4.2	Extension de l'entropie de Shannon aux capacités normalisées	35

2.4.3	Variance d'une capacité normalisée	37
2.4.4	Interprétation dans le contexte de l'agrégation par intégrale de Choquet	38
2.4.5	Extension de l'entropie de Shannon aux bi-capacités normalisées . .	39
2.5	Loi et moments de l'intégrale de Choquet dans le cas uniforme	41
2.5.1	Définitions : fonctions puissances tronquées et différences divisées .	41
2.5.2	Principaux résultats	41
2.5.3	Calcul de différences divisées de fonctions puissances tronquées . .	42
2.5.4	Exemple	43

2.1 Introduction

L'utilisation de mesures non additives et de leurs généralisations définies dans le chapitre précédent permet, dans de nombreux domaines, une représentation plus adéquate de la réalité. La flexibilité acquise en remplaçant la propriété d'additivité par d'autres propriétés moins contraignantes s'accompagne néanmoins d'une plus grande complexité en terme de représentation comme nous l'avons évoqué dans la Section 1.2.5. L'objectif de ce chapitre est de présenter les principaux indices numériques permettant de *synthétiser* l'information modélisée par les mesures non additives ou leurs extensions. Dans le contexte de l'aide multicritère à la décision fondée sur l'intégrale de Choquet, ces indices permettent de décrire les principales caractéristiques de l'agrégation, p. ex. en quantifiant l'*importance globale* des critères ou l'*interaction moyenne* entre critères.

La première section de ce chapitre est consacrée à la notion de *valeur* initialement introduite en théorie des jeux coopératifs mais trouvant des interprétations naturelles dans la plupart des domaines évoqués jusqu'à présent. Nous nous intéresserons plus particulièrement à son utilisation en aide multicritère à la décision fondée sur les intégrales non additives. Dans ce contexte, nous présenterons notamment les extensions proposées par Kojadinovic (2006c) permettant d'interpréter l'intégrale de Choquet bipolaire. La deuxième section porte sur la notion d'*indice d'interaction* proposée par Murofushi & Soneda (1993), puis généralisée par Roubens (1996) et Grabisch (1997b). Ce concept, récemment axiomatisé par Grabisch & Roubens (1999) et par Fujimoto, Kojadinovic & Marichal (2006), permet de déterminer le *type* et le *degré* d'interaction entre éléments d'un sous-ensemble de N . La troisième section a pour objet la notion d'*entropie*, récemment étendue aux capacités normalisées par Marichal & Roubens (2000b), et aux bi-capacités normalisées par Kojadinovic & Marichal (2006). Nous présenterons également des versions quadratiques de ces quantités (Kojadinovic 2006a) pouvant être utilisées pour l'identification de mesures non additives (cf. Chapitre 3). La dernière section est consacrée à la présentation des résultats obtenus par Marichal & Kojadinovic (2006) sur la loi et les moments de l'intégrale de Choquet dans le cas uniforme.

Il est important de noter que ce chapitre n'offre pas une vision exhaustive de l'interprétation de mesures non additives. Nous renvoyons le lecteur par exemple à (Grabisch 1997a, Marichal 2000b, Marichal 2004b, Marichal 2004a, Kojadinovic 2005, Grabisch 2005b) pour de plus amples détails sur l'interprétation de capacités dans le contexte de l'agrégation par intégrale de Choquet.

2.2 Notion de valeur

La notion de *valeur* est l'un des concepts les plus importants en théorie des jeux coopératifs notamment parce qu'elle apporte dans certains cas une solution au problème du *partage équitable* (voir p. ex. Moulin 2004). En théorie des votes, ce concept, appelé *indice de pouvoir*, est utilisé pour mesurer le *pouvoir* d'un votant (voir p. ex. Shapley & Shubik 1954, Banzhaf 1965). En aide multicritère à la décision fondée sur les intégrales non additives, la notion de valeur sert à quantifier l'importance d'un critère dans le processus d'agrégation, l'expression *indice d'importance* étant alors utilisée (voir p. ex. Marichal 2000b, Grabisch & Labreuche 2004).

Après avoir rappelé la définition de la valeur de Shapley (1953) ainsi que ses interprétations en théorie des jeux coopératifs et dans le contexte de l'agrégation par intégrale de Choquet, nous évoquerons brièvement la notion de *valeur généralisée* proposée par Marichal (2000c) et axiomatisée par Marichal, Kojadinovic & Fujimoto (2006). Nous présenterons ensuite les généralisations des indices d'importance proposées par Kojadinovic (2006c) permettant d'interpréter l'intégrale de Choquet bipolaire. Nous concluons cette section par une brève description des différentes extensions de la valeur de Shapley aux jeux bi-coopératifs récemment proposées dans la littérature (Labreuche & Grabisch 2006b).

2.2.1 Valeur de Shapley

Formellement, une *valeur* est une fonction de \mathcal{G}_N dans \mathbb{R}^n . Parmi les nombreuses valeurs définies en théorie des jeux coopératifs et en théorie des votes³, celle de Shapley (1953) est la plus connue et probablement la plus utilisée.

Définition 2.2.1 *La valeur de Shapley d'un jeu μ sur N est définie par*

$$\phi_\mu(i) := \sum_{T \subseteq N \setminus i} \gamma_t(n) [\mu(T \cup i) - \mu(T)], \quad \forall i \in N,$$

où

$$\gamma_t(n) := \frac{(n-t-1)!t!}{n!} \quad (t = 0, 1, \dots, n-1).$$

La valeur de Shapley d'un élément i par rapport à un jeu μ peut être vue comme une moyenne de la *contribution marginale* $\mu(T \cup i) - \mu(T)$ de i à un sous-ensemble T ne le contenant pas. En fonction de la représentation de Möbius de μ , la valeur de i s'écrit comme

$$\phi_\mu(i) = \sum_{T \subseteq N \setminus i} \frac{1}{t+1} m_\mu(T \cup i). \quad (2.1)$$

Interprétation en théorie des jeux coopératifs

Dans le contexte de la théorie des jeux coopératifs, la valeur de Shapley peut être vue comme une solution au problème du *partage équitable* (*fair share* en anglais, voir p. ex. Moulin 2004). Rappelons que, dans le cadre d'un jeu coopératif μ , $\mu(S)$ représente le gain que les joueurs de S

³Le site <http://powerslave.val.utu.fi/>, hébergé par l'université finlandaise de Turku, propose une bibliographie étendue sur les indices de pouvoir ainsi que des outils de calcul.

auraient à se partager si la coalition $S \subseteq N$ se formait. Il est sous-entendu que, finalement, seule la *grande coalition* N se formera. Le problème est alors de savoir comme partager le gain $\mu(N)$ parmi les n joueurs. Dans le cadre de la définition d'un mécanisme de partage, il paraît censé de tenir compte non seulement de $\mu(N)$ mais également des gains fictifs $\mu(S)$, $S \subsetneq N$. La valeur de Shapley propose alors un mécanisme de partage dans lequel la part du joueur i est $\phi_\mu(i)$. Cette solution est fondée sur des axiomes naturels, l'un des plus importants dans le contexte considéré étant l'axiome d'*efficacité* qui impose que

$$\sum_{i \in N} \phi_\mu(i) = \mu(N).$$

Interprétation dans le contexte de l'agrégation par intégrale de Choquet

Dans le contexte de l'agrégation par intégrale de Choquet, une interprétation probabiliste remarquable de la valeur de Shapley a été donnée par Marichal (1998).

Étant donné un vecteur $x \in \mathbb{R}^n$ et $a \in \mathbb{R}$, notons $(x|x_i = a)$ le vecteur de \mathbb{R}^n qui diffère de x seulement sur sa i -ème composante qui est égale à a . De plus, étant donnée une capacité normalisée μ sur N et $i \in N$, définissons

$$\Delta_i C_\mu(x) := C_\mu(x|x_i = 1) - C_\mu(x|x_i = 0).$$

Marichal a alors montré que

$$\int_{[0,1]^n} \Delta_i C_\mu(x) dx = \phi_\mu(i).$$

La valeur de Shapley de i par rapport à μ peut être ainsi vue comme l'espérance par rapport à la distribution uniforme sur $[0,1]^n$ de $\Delta_i C_\mu(x)$. Cette variation espérée peut être interprétée comme une mesure de l'*influence moyenne* du critère i sur le résultat de l'agrégation.

Il existe un lien encore plus évident qui conduit à une interprétation encore plus naturelle de la valeur de Shapley dans le contexte de l'agrégation par C_μ . Comme nous l'avons vu dans la Section 1.3.1, étant donné $x \in (\mathbb{R}^+)^n$, il existe $\sigma \in \Pi_N$ tel que $x \in \mathcal{K}_\sigma$ et ainsi

$$C_\mu(x) = \sum_{j \in N} \omega_\sigma^\mu(j) x_{\sigma(j)}.$$

La permutation σ correspond à la chaîne maximale c_σ de $(\mathcal{P}(N), \subseteq)$ induisant les poids de l'intégrale de Choquet, le poids du critère $i \in N$ étant $\omega_\sigma^\mu(\sigma^{-1}(i))$.

L'intégrale de Choquet pouvant être vue comme $n!$ moyennes arithmétiques pondérées, une façon naturelle de mesurer l'importance d'un critère $i \in N$ consiste simplement à calculer son poids moyen, c.-à-d.

$$\frac{1}{n!} \sum_{\sigma \in \Pi_N} \omega_\sigma^\mu(\sigma^{-1}(i)),$$

ce qui, en utilisant l'Eq. (1.4), peut être récrit comme

$$\frac{1}{n!} \sum_{\sigma \in \Pi_N} [\mu(\{i, \sigma(\sigma^{-1}(i) + 1), \dots, \sigma(n)\}) - \mu(\{\sigma(\sigma^{-1}(i) + 1), \dots, \sigma(n)\})].$$

Après changement de variable, nous obtenons

$$\frac{1}{n!} \sum_{\pi \in \Pi_N} [\mu(S_\pi^i \cup i) - \mu(S_\pi^i)], \quad (2.2)$$

où $S_\pi^i := \{j \in N \mid \pi(j) > \pi(i)\}$, c.-à-d. S_π^i est l'ensemble des successeurs de i dans l'ordre π .

L'expression donnée par l'Eq. (2.2) est une forme bien connue de la valeur de Shapley (1953) de i par rapport à μ . Il s'ensuit que, dans le contexte de l'agrégation par C_μ , $\phi_\mu(i)$ n'est rien d'autre que le poids moyen du critère i .

2.2.2 Valeurs généralisées

Le concept de *valeur* a été récemment généralisé par Marichal (2000c) afin de mesurer l'*influence globale* de chaque coalition dans un jeu. Cette généralisation résulte de l'étude de l'influence de variables sur les fonction booléennes et pseudo-booléennes (Bourgain, Kahn, Kalai, Katznelson & Linial 1992, Kahn, Kalai & Linial 1988) et de la recherche de schémas de vote robustes en théorie des votes (Ben-Or & Linial 1990).

Formellement, une valeur généralisée est une fonction de \mathcal{G}_N dans $\mathcal{P}(N)$. A titre d'exemple, nous donnons la définition de la valeur généralisée de Shapley proposée par Marichal (2000c).

Définition 2.2.2 *La valeur généralisée de Shapley d'un jeu μ sur N est définie par*

$$\Phi_\mu(S) := \sum_{T \subseteq N \setminus S} \gamma_t(n - s + 1)[\mu(T \cup S) - \mu(T)], \quad \forall S \subseteq N.$$

La valeur généralisée de Shapley d'une coalition S par rapport à un jeu μ peut être vue comme une moyenne de la *contribution marginale* $\mu(T \cup S) - \mu(T)$ de S à un sous-ensemble T de $N \setminus S$.

Les notions de *valeur probabiliste généralisée* et de *semivaleur généralisée* ont été récemment caractérisées par Marichal, Kojadinovic & Fujimoto (2006).

2.2.3 Indices d'importance dans le cas bipolaire

Avant de présenter les différentes extensions de la valeur de Shapley à des jeux bi-coopératifs proposées dans la littérature, nous introduisons les indices d'importance définis par Kojadinovic (2006c) permettant de mesurer l'influence d'un critère dans le contexte de l'agrégation par l'intégrale de Choquet bipolaire.

Une approche fondée sur la notion de poids d'un critère

L'interprétation de la valeur de Shapley par rapport une capacité normalisée donnée dans la Section 2.2.1 est basée sur le fait que l'intégrale de Choquet peut être regardée, dans ce cas, comme un ensemble de $n!$ moyennes arithmétiques pondérées. Comme nous l'avons vu dans la Section 1.4.4, il est possible d'avoir une vision similaire de l'intégrale de Choquet par rapport à une bi-capacité normalisée⁴ v sur N . En effet, étant donnée $x \in \mathbb{R}^n$, il existe $N^+ \subseteq N$ et $\sigma \in \Pi_N$

⁴Noter que, bien que les développements présentés dans cette section soient menés par rapport à des bi-capacités normalisées (la normalisation, comme évoqué dans la Section 1.4.4, est naturelle dans le contexte de l'agrégation), ils auraient pu être menés par rapport à des bi-capacités quelconques, la normalisation n'ayant pas de conséquences sur les définitions et propriétés présentées.

tels que $x \in \mathcal{K}_{N^+, \sigma}$ et, $\omega_{N^+, \sigma}^v(i)$ et $x_{\sigma(i)}$ étant de même signe pour tout $i \in N$,

$$C_v(x) = \sum_{j \in N} |\omega_{N^+, \sigma}^v(j)| x_{\sigma(j)}.$$

Le couple (N^+, σ) correspond à la chaîne maximale $c_{N^+, \sigma}$ de $(\mathcal{Q}(N), \subseteq)$ induisant les poids de l'intégrale de Choquet, le poids du critère $i \in N$ étant $|\omega_{N^+, \sigma}^v(\sigma^{-1}(i))|$. Ainsi, l'intégrale de Choquet par rapport à une bi-capacité peut être vue comme $2^n n!$ sommes pondérées.

Importance d'un critère

Afin de mesurer l'importance d'un critère durant l'agrégation par intégrale de Choquet bipolaire, il a été suggéré dans (Kojadinovic 2006c) de mesurer, comme précédemment, son poids moyen. Il faut néanmoins tenir compte du fait que, dans un contexte bipolaire, un critère peut être soit positivement soit négativement satisfait (cf. Section 1.3.5).

Définition 2.2.3 Soit v une bi-capacité normalisée sur N . Dans le cadre de l'agrégation par C_v , l'importance d'un critère $i \in N$ positivement satisfait est définie par

$$\phi_v^+(i) := \frac{1}{2^{n-1}} \sum_{i \in N^+ \subseteq N} \frac{1}{n!} \sum_{\sigma \in \Pi_N} |\omega_{N^+, \sigma}^v(\sigma^{-1}(i))|,$$

et l'importance d'un critère $i \in N$ négativement satisfait est définie par

$$\phi_v^-(i) := \frac{1}{2^{n-1}} \sum_{N^+ \subseteq N \setminus i} \frac{1}{n!} \sum_{\sigma \in \Pi_N} |\omega_{N^+, \sigma}^v(\sigma^{-1}(i))|.$$

Clairement, un indice d'importance global de i durant l'agrégation par C_v peut être obtenu en prenant la moyenne des deux indices définis ci-dessus, c.-à-d.

$$\phi_v(i) := \frac{1}{2} \phi_v^+(i) + \frac{1}{2} \phi_v^-(i) = \frac{1}{2^n} \sum_{N^+ \subseteq N} \frac{1}{n!} \sum_{\sigma \in \Pi_N} |\omega_{N^+, \sigma}^v(\sigma^{-1}(i))|.$$

Nous donnons ci-après les principales propriétés des indices d'importance définis ci-dessus. Des résultats supplémentaires peuvent être trouvés dans (Kojadinovic 2006c) et dans (Labreuche & Grabisch 2006b).

La proposition suivante donne la forme des indices d'importance directement en termes des coefficients de la bi-capacité.

Proposition 2.2.1 Pour toute bi-capacité normalisée v sur N et tout $i \in N$, nous avons

$$\phi_v^+(i) = \sum_{(S, T) \in \mathcal{Q}(N \setminus i)} \frac{\gamma_{s+t}(n)}{2^{s+t}} [v(S \cup i, T) - v(S, T)],$$

et

$$\phi_v^-(i) = \sum_{(S, T) \in \mathcal{Q}(N \setminus i)} \frac{\gamma_{s+t}(n)}{2^{s+t}} [v(S, T) - v(S, T \cup i)].$$

La propriété ci-après montre que la définition adoptée est consistante avec le fait que l'intégrale de Choquet par rapport à une bi-capacité est une extension de l'intégrale de Choquet par rapport à une capacité qui coïncide avec cette dernière lorsque la bi-capacité est asymétrique (cf. Eq. (1.7)).

Proposition 2.2.2 *Pour toute bi-capacité asymétrique normalisée $v_{\mu, \bar{\mu}}$ sur N et tout $i \in N$, nous avons*

$$\phi_v(i) = \frac{1}{2}\phi_{v_{\mu, \bar{\mu}}}^+(i) + \frac{1}{2}\phi_{v_{\mu, \bar{\mu}}}^-(i) = \phi_{\mu}(i).$$

La proposition suivante donne la forme des indices d'importance en fonction de la transformée de Möbius sur $(\mathcal{Q}(N), \subseteq)$.

Proposition 2.2.3 *Pour toute bi-capacité normalisée v sur N et tout $i \in N$, nous avons*

$$\phi_v^+(i) = \sum_{(S, T) \in \mathcal{Q}(N \setminus i)} \frac{1}{2^{s+t}} \frac{1}{s+t+1} m_v(S \cup i, T),$$

et

$$\phi_v^-(i) = - \sum_{(S, T) \in \mathcal{Q}(N \setminus i)} \frac{1}{2^{s+t}} \frac{1}{s+t+1} m_v(S, T \cup i).$$

Nous concluons cette section en donnant une interprétation probabiliste des indices d'importance proposés généralisant celle donnée par Marichal (1998) (cf. Section 2.2.1).

Rappelons que, étant donnée un vecteur $x \in \mathbb{R}^n$ et $a \in \mathbb{R}$, $(x|x_i = a)$ désigne le vecteur de \mathbb{R}^n qui diffère de x seulement sur sa i -ème composante laquelle vaut a . Étant donnée un bi-capacité normalisée v sur N et $i \in N$, définissons

$$\Delta_i^+ C_v(x) := C_v(x|x_i = 1) - C_v(x|x_i = 0),$$

et

$$\Delta_i^- C_v(x) := C_v(x|x_i = 0) - C_v(x|x_i = -1).$$

Proposition 2.2.4 *Pour toute bi-capacité normalisée v sur N et tout $i \in N$, nous avons*

$$\frac{1}{2^n} \int_{[-1, 1]^n} \Delta_i^+ C_v(x) dx = \frac{1}{2} \phi_v^+(i),$$

et

$$\frac{1}{2^n} \int_{[-1, 1]^n} \Delta_i^- C_v(x) dx = \frac{1}{2} \phi_v^-(i).$$

Les indices d'importance $\phi_v^+(i)$ et $\phi_v^-(i)$ peuvent être ainsi vus de façon équivalente comme les espérances par rapport à la distribution uniforme sur $[-1, 1]^n$ de $\Delta_i^+ C_v(x)$ et $\Delta_i^- C_v(x)$ respectivement. Comme dans le cas unipolaire, ces deux variations espérées peuvent être interprétées comme des mesures de l'influence moyenne du critère i sur le résultat de l'agrégation. De plus, nous avons clairement

$$\phi_v(i) = \frac{1}{2} \phi_v^+(i) + \frac{1}{2} \phi_v^-(i) = \frac{1}{2^n} \int_{[-1, 1]^n} \Delta_i C_v(x) dx,$$

où

$$\Delta_i C_v(x) := C_v(x|x_i = 1) - C_v(x|x_i = -1).$$

Importance relative d'un critère

Comme cela a été évoqué dans la Section 1.4.2, la distribution de poids impliquée dans le calcul de l'intégrale de Choquet dans le cas bipolaire n'est pas nécessairement une distribution de probabilité. Lorsqu'elle est normalisée comme dans l'Eq. (1.9), elle peut être regardée comme représentant le poids relatif des critères (cf. Eq. (1.8)).

En procédant comme précédemment, l'importance relative d'un critère $i \in N$ positivement satisfait a été définie dans (Kojadinovic 2006c) par

$$\psi_v^+(i) := \frac{1}{2^{n-1}} \sum_{i \in N^+ \subseteq N} \frac{1}{n!} \sum_{\sigma \in \Pi_N} p_{N^+, \sigma}^v(\sigma^{-1}(i)),$$

et l'importance relative d'un critère $i \in N$ négativement satisfait par

$$\psi_v^-(i) := \frac{1}{2^{n-1}} \sum_{N^+ \subseteq N \setminus i} \frac{1}{n!} \sum_{\sigma \in \Pi_N} p_{N^+, \sigma}^v(\sigma^{-1}(i)).$$

Comme précédemment, les deux indices peuvent être combinés en un indice global en prenant leur moyenne :

$$\psi_v(i) := \frac{1}{2} \psi_v^+(i) + \frac{1}{2} \psi_v^-(i), \quad \forall i \in N.$$

Les indices d'importance relative proposés coïncident avec les indices d'importance définis précédemment dès que la bi-capacité est asymétrique, car, dans ce cas, le poids d'un critère et son poids relatif sont la même chose.

Proposition 2.2.5 *Pour toute bi-capacité asymétrique normalisée $v_{\mu, \bar{\mu}}$ sur N et pour tout $i \in N$, nous avons*

$$\psi_{v_{\mu, \bar{\mu}}}^+(i) = \phi_{v_{\mu, \bar{\mu}}}^+(i) \quad \text{et} \quad \psi_{v_{\mu, \bar{\mu}}}^-(i) = \phi_{v_{\mu, \bar{\mu}}}^-(i).$$

Corollaire 2.2.1 *Pour toute bi-capacité asymétrique normalisée $v_{\mu, \bar{\mu}}$ sur N et pour tout $i \in N$, nous avons*

$$\psi_{v_{\mu, \bar{\mu}}}(i) = \frac{1}{2} \psi_{v_{\mu, \bar{\mu}}}^+(i) + \frac{1}{2} \psi_{v_{\mu, \bar{\mu}}}^-(i) = \phi_{\mu}(i).$$

Nous donnons ci-après une dernière propriété qui peut être regardée comme l'analogue de la propriété d'efficacité satisfaite par la valeur de Shapley.

Proposition 2.2.6 *Pour toute bi-capacité v sur N et pour tout $i \in N$, nous avons*

$$\sum_{i \in N} \psi_v(i) = 1.$$

Notons que dans le cadre de l'application de modèles décisionnels fondés sur l'intégrale de Choquet bipolaire, les indices d'importance relative devraient a priori être plus faciles à interpréter pour un décideur que les indices définis dans la section précédente.

2.2.4 Valeur de Shapley par rapport à un jeu bi-coopératif

Plusieurs extensions de la valeur de Shapley à des jeux bi-coopératifs ont été récemment proposées dans la littérature. La première, due à Felsenthal & Machover (1997), est très spécifique aux jeux de votes ternaires et ne semble pas adaptée à la théorie des jeux bi-coopératifs. Une deuxième proposition, liée à la définition de la notion de bi-capacité et à l'intégrale de Choquet bipolaire, est due à Grabisch & Labreuche (2005a). L'inconvénient principal de l'indice proposé est qu'il n'a pas d'interprétation naturelle, ni dans le contexte de l'agrégation par intégrale de Choquet, ni par rapport à un jeu bi-coopératif. Une troisième définition a été proposée par Bilbao, Fernández, Jiménez & López (2006). L'approche des auteurs a grosso modo consisté à transposer les propriétés de la valeur de Shapley en termes de chaînes maximales de $(\mathcal{P}(N), \subseteq)$ sur $(\mathcal{Q}(N), \sqsubseteq)$ (cf. Section 1.4.2). La proposition la plus récente, et la plus convaincante dans le contexte des jeux bi-coopératifs, est due à Labreuche & Grabisch (2006b). L'approche proposée par les auteurs s'appuie sur des axiomes naturels dans le contexte de jeux bi-coopératifs. La valeur résultante coïncide au signe près avec les indices d'importance définis dans la section précédente.

2.3 Indices d'interaction

Le fait que, dans le cas général, la valeur d'un élément $i \in N$ relativement à un jeu coopératif μ ne soit pas égale au coefficient $\mu(i)$ montre l'existence de phénomènes d'interaction entre éléments de N . Afin de décrire ces phénomènes, Murofushi & Soneda (1993), puis Grabisch (1997b) et Roubens (1996) ont proposé des indices, appelés *indices d'interaction*, permettant de déterminer le *type* et le *degré* d'interaction entre éléments d'un sous-ensemble de N .

Nous présenterons tout d'abord les indices d'interaction de Shapley proposés par Murofushi & Soneda (1993) et Grabisch (1997b). Nous donnerons ensuite leur interprétation dans le contexte de l'agrégation par intégrale de Choquet. Enfin, nous mentionnerons brièvement l'extension de ces indices proposée par Kojadinovic (2006c) permettant d'interpréter le comportement de l'intégrale de Choquet bipolaire.

2.3.1 Notion d'interaction

Afin d'approcher la notion d'interaction de façon intuitive, considérons deux éléments i et j de N tels que l'importance de la paire ij soit strictement supérieure à la somme des importances de i et de j . L'inégalité

$$\mu(ij) > \mu(i) + \mu(j)$$

semble alors traduire une *synergie positive* entre i et j et il apparaît naturel de dire que i et j *interagissent* de façon *positive* ou encore *complémentaire*. De façon analogue, l'inégalité

$$\mu(ij) < \mu(i) + \mu(j)$$

suggère une interaction *négative* ou *substitutive* entre i et j . Enfin, si

$$\mu(ij) = \mu(i) + \mu(j),$$

l'intuition nous pousse à considérer que les éléments i et j *n'interagissent pas* ou encore qu'ils sont *indépendants*.

L'*interaction* entre i et j semble ainsi clairement dépendre de la différence $\mu(ij) - \mu(i) - \mu(j)$. Dans la Section 2.2.1, nous avons néanmoins vu que l'*importance globale* d'un élément $i \in N$ dépend non seulement du coefficient $\mu(i)$ mais également de la contribution marginale de i à toutes les coalitions ne le contenant pas, c.-à-d. de la famille $(\mu(T \cup i) - \mu(T))_{T \subseteq N \setminus i}$. De la même façon, un coefficient d'interaction entre deux éléments i et j devrait prendre en compte non seulement les coefficients $\mu(i)$, $\mu(j)$ et $\mu(ij)$ mais également tous les coefficients de la forme $\mu(T \cup i)$, $\mu(T \cup j)$ et $\mu(T \cup ij)$, $T \subseteq N \setminus ij$.

2.3.2 Indices d'interaction de Shapley par rapport à un jeu coopératif

Murofushi & Soneda (1993) ont proposé de mesurer l'interaction moyenne entre deux éléments i et j par rapport à un jeu μ sur N par le biais de l'*indice d'interaction* suivant :

$$I_\mu(ij) := \sum_{T \subseteq N \setminus ij} \gamma_t(n-1) [\mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)].$$

Remarquons que, pour un sous-ensemble T ne contenant pas i et j , l'expression

$$\mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)$$

peut être vue comme la différence entre les contributions marginales $\mu(T \cup ij) - \mu(T \cup i)$ et $\mu(T \cup j) - \mu(T)$. À l'instar de Grabisch, Marichal & Roubens (2000), nous appellerons cette expression l'*interaction marginale entre i et j en présence de T* . En effet, il semble naturel de considérer que si

$$\mu(T \cup ij) - \mu(T \cup i) > \mu(T \cup j) - \mu(T) \text{ (resp. } < \text{)},$$

i et j interagissent *positivement* (resp. *négativement*) en présence de T .

L'indice proposé par Murofushi & Soneda peut alors être interprété comme une moyenne pondérée de l'interaction marginale entre i et j en présence de T , toutes les coalitions T ne contenant pas i et j étant prises en compte. Une propriété importante est que $I_\mu(ij) \in [-1, 1]$ pour tout $ij \subseteq N$ (Grabisch 1997b), la valeur 1 (resp. -1) correspondant à une complémentarité (resp. substitutivité) maximale entre i et j .

Grabisch (1997b) a étendu l'indice ci-dessus aux coalitions contenant plus de deux joueurs.

Définition 2.3.1 *L'indice d'interaction de Shapley d'une capacité μ sur N est défini par*

$$I_\mu(S) := \sum_{T \subseteq N \setminus S} \gamma_t(n-s+1) \sum_{L \subseteq S} (-1)^{s-l} \mu(L \cup T), \quad \forall S \subseteq N.$$

Cet indice est une extension de la valeur de Shapley dans le sens où, pour tout jeu μ sur N , $I_\mu(i)$ et $\phi_\mu(i)$ coïncident pour tout $i \in N$. Il a été caractérisé dans le contexte de la théorie des jeux coopératifs par Grabisch & Roubens (1999). Une caractérisation alternative a été proposée par Fujimoto, Kojadinovic & Marichal (2006), qui ont également axiomatisé les classes d'indices d'interaction *probabilistes* et *cardinal-probabilistes*.

Pour tout $S \subseteq N$ tel que $s \geq 2$, une valeur strictement positive (resp. négative) de $I_\mu(S)$ peut être interprétée comme une interaction complémentaire (resp. substitutive) *simultanée* entre tous les critères de S *en moyenne* (Fujimoto, Kojadinovic & Marichal 2006).

En fonction de la représentation de Möbius de μ , l'indice d'interaction de Shapley est donné par

$$I_\mu(S) = \sum_{T \supseteq S} \frac{1}{t - s + 1} m_\mu(T), \quad \forall S \subseteq N. \quad (2.3)$$

Interprétation dans le contexte de l'agrégation par intégrale de Choquet

Nous présentons l'interprétation de l'indice d'interaction de Shapley donnée dans (Kojadinovic 2006c). Dans le contexte de l'agrégation par l'intégrale de Choquet par rapport à une capacité normalisée μ sur N , l'indice d'interaction de Shapley de $ij \subseteq N$ peut s'écrire comme la différence de poids moyens suivante :

$$I_\mu(ij) = \frac{1}{(n-1)!} \sum_{\substack{\sigma \in \Pi_N \\ \sigma(n)=j}} \omega_\sigma^\mu(\sigma^{-1}(i)) - \frac{1}{(n-1)!} \sum_{\substack{\sigma \in \Pi_N \\ \sigma(1)=j}} \omega_\sigma^\mu(\sigma^{-1}(i)).$$

L'expression

$$\frac{1}{(n-1)!} \sum_{\substack{\sigma \in \Pi_N \\ \sigma(n)=j}} \omega_\sigma^\mu(\sigma^{-1}(i)) \quad \left(\text{resp. } \frac{1}{(n-1)!} \sum_{\substack{\sigma \in \Pi_N \\ \sigma(1)=j}} \omega_\sigma^\mu(\sigma^{-1}(i)) \right)$$

peut être interprétée comme le poids moyen du critère i lorsque j est le critère le mieux satisfait (resp. le moins bien satisfait). La quantité $I_\mu(ij)$ peut ainsi être regardée comme la variation du poids moyen i lorsque j passe de critère le moins bien satisfait au critère le mieux satisfait. Une valeur négative de $I_\mu(ij)$ correspond à la situation où le poids moyen de i lorsque j est le critère le mieux satisfait est inférieur au poids moyen de i lorsque j est le critère le moins bien satisfait. Ce comportement traduit une tendance à compenser les évaluations faibles sur j en augmentant le poids de i , i.e., un comportement *tolérant* ou *disjonctif*. Une valeur positive correspond à la situation opposée : le passage de j de critère le moins bien satisfait au critère le mieux satisfait conduit à une diminution du poids de i , ce qui indique un comportement *intolérant* ou *conjonctif*. Par symétrie, une interprétation similaire par rapport au poids de j est bien évidemment possible.

Pour trois critères $ijk \subseteq N$, un raisonnement similaire conduit à une interprétation de $I_\mu(ijk)$ comme une différence de second ordre de poids moyens. Plus généralement, $I_\mu(S)$, $S \subseteq N$, $s \geq 2$, peut être vu comme une différence d'ordre $s - 1$ de poids moyens.

2.3.3 Extension au cas bipolaire

En partant des indices d'importance définis à la Section 2.2.3 et en utilisant des propriétés de récursivité généralisant celles satisfaites par les indices d'interaction dans le cas unipolaire (utilisées également dans Grabisch & Labreuche 2005a, Lange & Grabisch 2005), des indices d'interaction permettant d'interpréter l'intégrale de Choquet bipolaire ont été proposés par Kojadinovic (2006c). Comme dans le cas unipolaire, ces indices peuvent s'interpréter comme des variations de poids moyens.

2.4 Mesures d'uniformité

Les capacités pouvant être vues comme des généralisations de mesures de probabilité, il semble naturel de s'interroger sur la possibilité d'étendre les mesures d'uniformité usuelles, en particulier les mesures d'*entropie*, à ces fonctions d'ensemble non additives. Dans le cadre de la théorie des fonctions de croyance et de la théorie des possibilités, différentes mesures d'*incertitude*, spécifiques à ces domaines, ont été définies (Klir & Yuan 1995, Klir & Wierman 1999, Dubois & Prade 1987, Dubois & Ramer 1993, Klir & Smith 2001). Récemment, Marichal & Roubens (2000b) ont proposé un indice qui peut être interprété comme une extension "directe" de l'entropie de Shannon aux capacités normalisées. Cette quantité a été récemment axiomatisée en tant que mesure d'uniformité par Kojadinovic, Marichal & Roubens (2005) et, dans un cadre plus général, par Honda & Grabisch (2006).

L'objectif de cette section est, dans un premier temps, de rappeler la définition et les propriétés de l'entropie proposée par Marichal & Roubens (2000b). Dans un deuxième temps, nous présenterons les mesures d'uniformité quadratiques définies et étudiées par Kojadinovic (2006a) qui seront utilisées pour l'identification de capacités dans le Chapitre 3. Enfin, l'extension de l'entropie de Shannon aux bi-capacités proposée par Kojadinovic & Marichal (2006) sera présentée et ses principales propriétés seront énoncées.

Dans la suite de cette section, nous emploierons indifféremment les expressions *mesure d'entropie* et *mesure d'uniformité*. En revanche, regarder les différentes mesures qui vont être présentées comme des mesures d'*incertitude* n'a pas beaucoup de sens dans le cadre général considéré. Nous verrons néanmoins qu'une interprétation en tant que mesures d'*information utilisée* peut en être donnée dans le contexte de l'agrégation par intégrale de Choquet.

2.4.1 Notion d'entropie

Le concept fondamental d'*entropie d'une distribution de probabilité* a été initialement proposé par Shannon (1948). L'entropie de Shannon d'une distribution de probabilité p sur N est définie par

$$H_S(p) := \sum_{i \in N} h[p(i)]$$

où

$$h(x) := \begin{cases} -x \ln x, & \text{si } x > 0, \\ 0, & \text{si } x = 0, \end{cases}$$

La quantité $H_S(p)$ est toujours non négative et vaut zéro si et seulement si p est une masse de Dirac (propriété dite de *decisivité*). En tant que fonction de p , H_S est strictement concave. De plus, elle atteint sa valeur maximale ($\ln n$) si et seulement si p est uniforme (propriété dite de *maximalité*).

Dans un contexte non probabiliste, $H_S(p)$ peut être vue comme une mesure de l'uniformité ou encore de la régularité de p . Dans un contexte probabiliste, lorsque p sert à modéliser un système stochastique discret à n états, l'entropie de p s'interprète naturellement comme une mesure de l'incertitude liée à un état futur du système.

Plusieurs caractérisations axiomatiques de l'entropie de Shannon ont été proposées dans la littérature (voir p. ex. Khinchin 1957, Aczél & Daróczy 1975, Ebanks, Sahoo & Sander 1997, Jaynes 2003), la plus connue étant probablement celle de Faddeev (1957).

Une généralisation paramétrique bien connue de l'entropie de Shannon est l'entropie de Havrda & Charvat (1967) d'ordre β définie, pour tout réel strictement positif β et toute distribution de probabilité p sur N , par

$$H_{HC}^\beta(p) := \begin{cases} \frac{1}{1-\beta} \left[\sum_{i \in N} p(i)^\beta - 1 \right], & \beta \neq 1, \\ H_S(p), & \beta = 1. \end{cases} \quad (2.4)$$

Comme l'entropie de Shannon, l'entropie de Havrda & Charvat d'ordre β est une fonction strictement concave de la distribution de probabilité et satisfait les propriétés de décisivité et de maximalité (avec l'exception que sa valeur maximale n'est $\ln n$ que lorsque $\beta = 1$).

La caractérisation axiomatique de l'entropie H_{HC}^β proposée par Havrda & Charvat (1967) est très similaire à celle de l'entropie de Shannon proposée par Faddeev (1957). La seule différence provient de la forme de l'axiome de récursivité.

Notons que de nombreuses autres généralisations de l'entropie de Shannon ont été proposées dans la littérature en affaiblissant les axiomes sur lesquels elle repose. La volonté de caractériser ces mesures d'entropies généralisées a conduit à la redécouverte de mesures d'*uniformité* et de *diversité* introduites au début du XX^e siècle (voir p. ex. Morales, Pardo & Vajda 1996). Le nombre de mesures d'entropie proposées dans la littérature est en fait si vaste (Esteban & Morales 1995) qu'il est difficile de croire que chacune de ces mesures ait été définie en réponse à un problème particulier.

2.4.2 Extension de l'entropie de Shannon aux capacités normalisées

Marichal & Roubens (2000b) ont proposé une extension naturelle de l'entropie de Shannon aux capacités normalisées (voir aussi Marichal 1998, Marichal 2002b). D'autres définitions moins convaincantes sont dues à Yager (2000) et à Dukhovny (2002).

Définition 2.4.1 *L'extension de l'entropie de Shannon à une capacité normalisée μ sur N est définie par*

$$H_{MR}(\mu) := \sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) h[\mu(S \cup i) - \mu(S)].$$

L'entropie H_{MR} a été récemment axiomatisée en tant que mesure d'uniformité par Kojadinovic, Marichal & Roubens (2005) et, dans un cadre plus général, par Honda & Grabisch (2006).

Une propriété fondamentale de H_{MR} est qu'elle peut être réécrite en fonction des chaînes maximales de $(\mathcal{P}(N), \subseteq)$ (Kojadinovic, Marichal & Roubens 2005) (voir aussi la Section 1.2.6). Pour toute capacité normalisée μ sur N , nous avons

$$H_{MR}(\mu) = \frac{1}{n!} \sum_{c \in \mathcal{M}(\mathcal{P}(N), \subseteq)} H_S(\omega_c^\mu),$$

où de façon équivalente,

$$H_{MR}(\mu) = \frac{1}{n!} \sum_{\sigma \in \Pi_N} H_S(\omega_\sigma^\mu). \quad (2.5)$$

La quantité $H_{MR}(\mu)$ peut ainsi être simplement vue comme une moyenne sur $\mathcal{M}_{(\mathcal{P}(N), \subseteq)}$ des entropies de Shannon des distributions de probabilité ω_c^μ . De façon équivalente, $H_{MR}(\mu)$ peut être interprétée comme une mesure de la *régularité moyenne* de la monotonie de μ le long des chaînes maximales de $(\mathcal{P}(N), \subseteq)$. Afin de garder à l'esprit que H_{MR} est une moyenne d'entropies de Shannon, nous la noterons \overline{H}_S dans la suite.

Les propriétés les plus importantes de $H_{MR} = \overline{H}_S$ sont (Marichal 2002b, Kojadinovic, Marichal & Roubens 2005) :

1. **Entropie d'une capacité additive.** Pour toute capacité additive $\mu \in \mathcal{C}_N^*$, nous avons

$$\overline{H}_S(\mu) = H_S(p),$$

où p est la distribution de probabilité sur N définie par $p(i) = \mu(i)$ pour tout $i \in N$.

2. **Entropie d'une capacité cardinale.** Pour toute capacité cardinale $\mu \in \mathcal{C}_N^*$, nous avons

$$\overline{H}_S(\mu) = H_S(p^\mu),$$

où p^μ est la distribution de probabilité sur N définie $p^\mu(i) = \mu(\{i, \dots, n\}) - \mu(\{i+1, \dots, n\})$ pour tout $i \in N$.

3. **Symétrie.** Pour toute capacité $\mu \in \mathcal{C}_N^*$ et toute permutation $\pi \in \Pi_N$, nous avons

$$\overline{H}_S(\mu \circ \pi) = \overline{H}_S(\mu).$$

4. **Expansibilité.** Soit μ une capacité normalisée sur N . Si $k \in N$ est un *élément nul* pour μ , c.-à-d. $\mu(T \cup k) = \mu(T)$ pour $T \subseteq N \setminus k$, alors

$$\overline{H}_S(\mu) = \overline{H}_S(\mu^{N \setminus k}),$$

où $\mu^{N \setminus k}$ désigne la restriction de μ à $\mathcal{P}(N \setminus k)$.

5. **Décisivité.** Pour toute capacité $\mu \in \mathcal{C}_N^*$,

$$\overline{H}_S(\mu) \geq 0.$$

De plus, $\overline{H}_S(\mu) = 0$ si et seulement si μ est une capacité binaire, c.-à-d. telle que $\mu(T) \in \{0, 1\}$ pour tout $T \subseteq N$.

6. **Maximalité.** Pour toute capacité $\mu \in \mathcal{C}_N^*$, nous avons

$$\overline{H}_S(\mu) \leq \ln n,$$

avec égalité si et seulement si μ est la capacité uniforme μ^* sur N .

7. **Monotonie croissante vers μ^* .** Soit μ une capacité normalisée sur N telle que $\mu \neq \mu^*$ et, pour tout $\lambda \in [0, 1]$, soit μ_λ la capacité normalisée sur N définie par $\mu_\lambda := \mu + \lambda(\mu^* - \mu)$. Alors, pour tout $0 \leq \lambda_1 < \lambda_2 \leq 1$, nous avons

$$\overline{H}_S(\mu_{\lambda_1}) < \overline{H}_S(\mu_{\lambda_2}).$$

8. **Stricte concavité.** Pour toute paire de capacités $\mu_1, \mu_2 \in \mathcal{C}_N^*$ et tout $\lambda \in]0, 1[$, nous avons

$$\overline{H}_S(\lambda \mu_1 + (1 - \lambda) \mu_2) > \lambda \overline{H}_S(\mu_1) + (1 - \lambda) \overline{H}_S(\mu_2).$$

La dernière propriété implique que maximiser \overline{H}_S sur un sous-ensemble convexe de \mathcal{C}_N^* conduit toujours à un optimum unique.

En raisonnant de façon similaire sur les distributions de probabilités induites par les chaînes maximales de $(\mathcal{P}(N), \subseteq)$ (cf. Section 1.2.6), l'entropie de Havrda & Charvat a été également étendue aux capacités normalisées (Kojadinovic 2006a).

Définition 2.4.2 *L'extension de l'entropie de Havrda & Charvat à une capacité normalisée μ sur N est définie par*

$$\overline{H}_{HC}^\beta(\mu) := \frac{1}{1-\beta} \left[\sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) [\mu(S \cup i) - \mu(S)]^\beta - 1 \right], \quad \beta > 0, \beta \neq 1. \quad (2.6)$$

À partir de la caractérisation et des propriétés de l'entropie probabiliste de Havrda & Charvat définie par l'Eq. (2.4), il peut être vérifié que \overline{H}_{HC}^β satisfait les huit propriétés données ci-dessus pour \overline{H}_S (moyennant une modification de la valeur maximale donnée dans la propriété de maximalité).

2.4.3 Variance d'une capacité normalisée

Une autre façon immédiate de mesurer l'uniformité d'une distribution de probabilité consiste à calculer sa *variance*. En procédant comme précédemment, ce concept a été étendu aux capacités normalisées par Kojadinovic (2006a).

Définition 2.4.3 *La variance d'une capacité normalisée μ sur N est définie par*

$$\overline{V}(\mu) := \frac{1}{n} \sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) \left(\mu(S \cup i) - \mu(S) - \frac{1}{n} \right)^2. \quad (2.7)$$

Il est facile de vérifier que, pour toute capacité $\mu \in \mathcal{C}_N^*$, l'entropie de Havrda & Charvat d'ordre 2 et la variance sont liées par l'équation suivante :

$$\overline{H}_{HC}^2(\mu) = \frac{n-1}{n} - n\overline{V}(\mu).$$

En utilisant par exemple l'égalité précédente, il peut être vérifié que \overline{V} satisfait entre autres les propriétés suivantes :

1. **Minimalité.** Pour toute capacité $\mu \in \mathcal{C}_N^*$, nous avons

$$\overline{V}(\mu) \geq 0.$$

avec égalité si et seulement si μ est la capacité uniforme μ^* sur N .

2. **Décisivité.** Pour toute capacité $\mu \in \mathcal{C}_N^*$,

$$\overline{V}(\mu) \leq \frac{n-1}{n^2}$$

avec égalité si et seulement si μ est une capacité binaire, c.-à-d. telle que $\mu(T) \in \{0, 1\}$ pour tout $T \subseteq N$.

3. **Monotonie décroissante vers μ^* .** Soit μ une capacité normalisée sur N telle que $\mu \neq \mu^*$ et, pour tout $\lambda \in [0, 1]$, soit μ_λ la capacité normalisée sur N définie par $\mu_\lambda := \mu + \lambda(\mu^* - \mu)$. Alors, pour tout $0 \leq \lambda_1 < \lambda_2 \leq 1$, nous avons

$$\bar{V}(\mu_{\lambda_1}) > \bar{V}(\mu_{\lambda_2}).$$

4. **Stricte convexité.** Pour toute paire de capacités $\mu_1, \mu_2 \in \mathcal{C}_N^*$ et tout $\lambda \in]0, 1[$, nous avons

$$\bar{V}(\lambda \mu_1 + (1 - \lambda) \mu_2) < \lambda \bar{V}(\mu_1) + (1 - \lambda) \bar{V}(\mu_2).$$

Le caractère quadratique de \bar{V} et cette dernière propriété sont à l'origine de la méthode d'identification de capacités proposée par Kojadinovic (2006a), dite du *minimum de variance*, que nous présenterons dans la Section 3.3.3.

2.4.4 Interprétation dans le contexte de l'agrégation par intégrale de Choquet

Dans le contexte de l'agrégation par intégrale de Choquet, les mesures d'uniformité présentées précédemment peuvent être utilisées pour quantifier le *degré d'utilisation* moyen des évaluations partielles lors de l'agrégation.

Supposons que, dans le cadre d'un problème d'aide multicritère à la décision, les évaluations partielles soient données sur l'échelle unipolaire $[0, 1]$. Comme nous l'avons vu dans la Section 1.3.1, pour un vecteur d'évaluations partielles $x \in [0, 1]^n$, il existe $\sigma \in \Pi_N$ tel que $x \in \mathcal{O}_\sigma := \mathcal{K}_\sigma \cap [0, 1]^n$ et

$$C_\mu(x) = \sum_{i \in N} \omega_\sigma^\mu(i) x_{\sigma(i)}.$$

Dans ce contexte, $H_S(\omega_\sigma^\mu)$ est une mesure de l'uniformité de la distribution ω_σ^μ ou encore de la régularité de l'accroissement de μ le long de la chaîne c_σ de $(\mathcal{P}(N), \subseteq)$. De façon équivalente, $H_S(\omega_\sigma^\mu)$ mesure le *degré d'utilisation* du vecteur x lors du calcul de la valeur agrégée $C_\mu(x)$. En effet, si $H_S(\omega_\sigma^\mu)$ est proche de $\ln n$, cela implique que la distribution de probabilité ω_σ^μ est à peu près uniforme et donc que tous les scores partiels x_i ont sensiblement la même influence sur $C_\mu(x)$ qui sera ainsi proche de la moyenne arithmétique des x_i . En revanche, si $H_S(\omega_\sigma^\mu)$ est proche de zéro, alors nécessairement un des $\omega_\sigma^\mu(i)$ est proche de 1 et $C_\mu(x)$ sera proche du score partiel correspondant.

En partant de l'Eq. (2.5) et en utilisant le fait que

$$\int_{x \in \mathcal{O}_\sigma} dx = 1/n!$$

l'entropie $\bar{H}_S(\mu)$ peut se réécrire comme

$$\bar{H}_S(\mu) = \sum_{\sigma \in \Pi_N} \int_{x \in \mathcal{O}_\sigma} H_S(\omega_\sigma^\mu) dx = \int_{[0, 1]^n} H_S(\omega_{\sigma_x}^\mu) dx,$$

où σ_x est une permutation de Π_N telle que $x \in \mathcal{O}_{\sigma_x}$. Ainsi, $\bar{H}_S(\mu)$ peut être interprétée comme la *valeur moyenne* sur l'ensemble des $x \in [0, 1]^n$ du *degré d'utilisation* des évaluations partielles x_1, \dots, x_n lors du calcul de $C_\mu(x)$ (Kojadinovic, Marichal & Roubens 2005).

Notons qu'en pratique, il est plus approprié de calculer la version normalisée de $\overline{H}_S(\mu)$, définie par $\overline{H}_S^*(\mu) := \overline{H}_S(\mu) / \ln n$. Une valeur de $\overline{H}_S^*(\mu)$ proche de 1 peut alors être interprétée comme une agrégation "neutre", dans le sens où l'intégrale de Choquet se confond alors presque avec la moyenne arithmétique simple. À l'opposée, une valeur de $\overline{H}_S^*(\mu)$ égale à 0 correspond à une agrégation très extrême, dictatoriale, car seule une évaluation partielle sera utilisée lors du calcul de la valeur agrégée.

Les quantités $\overline{H}_{HC}^\beta(\mu)$ et $\overline{V}(\mu)$ peuvent être normalisées et interprétées de façon similaire.

2.4.5 Extension de l'entropie de Shannon aux bi-capacités normalisées

La définition suivante a été proposée par Kojadinovic & Marichal (2006) afin de généraliser, dans l'esprit des sections précédentes, l'entropie de Shannon aux bi-capacités normalisées.

Définition 2.4.4 *L'extension de l'entropie de Shannon à une bi-capacité normalisée v sur N est définie par*

$$\overline{\overline{H}}_S(v) := \frac{1}{2^n} \sum_{N^+ \subseteq N} \frac{1}{n!} \sum_{\sigma \in \Pi_N} H_S(p_{\sigma, N^+}^v), \quad (2.8)$$

où p_{σ, N^+}^v , $\sigma \in \Pi_N$, $N^+ \subseteq N$, est la distribution de probabilité définie par l'Eq. (1.9).

Comme dans le cas des capacités normalisées, l'entropie de Shannon d'une bi-capacité normalisée sur N n'est rien d'autre que la moyenne des entropies de Shannon des distributions de probabilité induites par les chaînes maximales de $(\mathcal{Q}(N), \subseteq)$.

Dans le contexte de l'agrégation par l'intégrale de Choquet par rapport à une bi-capacité normalisée v sur N , $\overline{\overline{H}}_S(v)$ peut être vue comme une mesure de la valeur moyenne sur l'ensemble des $x \in [-1, 1]^n$ du *degré d'utilisation* des évaluations partielles x_1, \dots, x_n lors du calcul de la valeur agrégée $C_v(x)$. En effet, pour $x \in [-1, 1]^n$, il existe $N^+ \subseteq N$ et $\sigma \in \Pi_N$ tels que $x \in \mathcal{O}_{\sigma, N^+} := \mathcal{K}_{\sigma, N^+} \cap [-1, 1]^n$ et, comme nous pouvons le voir à partir de l'Eq. (1.8), $C_v(x)$ est alors proportionnel à $\sum_{i \in N} p_{\sigma, N^+}^v(i) x_{\sigma(i)}$.

En partant de l'Eq. (2.8) et en utilisant le fait que

$$\int_{x \in \mathcal{O}_{\sigma, N^+}} dx = 1/n!$$

l'entropie $\overline{\overline{H}}_S(v)$ peut se récrire comme

$$\begin{aligned} \overline{\overline{H}}_S(\mu) &= \frac{1}{2^n} \sum_{N^+ \subseteq N} \sum_{\sigma \in \Pi_N} \int_{x \in \mathcal{O}_{\sigma, N^+}} H_S(p_{\sigma, N^+}^v) dx \\ &= \frac{1}{2^n} \int_{[-1, 1]^n} H_S(p_{\sigma_x, N_x^+}^v) dx, \end{aligned}$$

où $N_x^+ \subseteq N$ et $\sigma_x \in \Pi_N$ sont définis tels que $x \in \mathcal{O}_{\sigma_x, N_x^+}$. Nous voyons ainsi que $\overline{\overline{H}}_S(v)$ mesure la valeur moyenne sur l'ensemble des $x \in [-1, 1]^n$ du degré d'utilisation des évaluations partielles x_1, \dots, x_n lors du calcul de $C_v(x)$.

La quantité $\overline{\overline{H}}_S$ satisfait plusieurs propriétés qui peuvent être considérées comme naturelles pour une mesure d'entropie. Ces propriétés, qui généralisent celles satisfaites par \overline{H}_S , ont été prouvées dans (Kojadinovic & Marichal 2006).

1. **Entropie d'une bi-capacité asymétrique.** Pour toute bi-capacité asymétrique normalisée $v_{\mu, \bar{\mu}}$ sur N , nous avons

$$\overline{\overline{H}}_S(v_{\mu, \bar{\mu}}) = \overline{H}_S(\mu).$$

Notons que cette propriété est complètement en accord avec l'Eq. (1.7).

2. **Entropie d'une bi-capacité additive.** Pour toute bi-capacité additive normalisée v_{μ_1, μ_2} sur N , nous avons

$$\overline{\overline{H}}_S(v_{\mu_1, \mu_2}) = \frac{1}{2^n} \sum_{N^+ \subseteq N} \sum_{i \in N} h \left[\frac{\mu_1(i \cap N^+) + \mu_2(i \cap N^-)}{\mu_1(N^+) + \mu_2(N^-)} \right]$$

3. **Entropie d'une bi-capacité additive symétrique/asymétrique.** Pour toute bi-capacité additive symétrique/asymétrique normalisée $v_{\mu, \mu}$ sur N , nous avons

$$\overline{\overline{H}}_S(v_{\mu, \mu}) = H_S(p),$$

où p est la distribution de probabilité sur N définie par $p(i) := \mu(i)$ pour tout $i \in N$.

4. **Symétrie.** Pour toute bi-capacité $v \in \mathcal{BC}_N^*$ et toute permutation π sur N , nous avons

$$\overline{\overline{H}}_S(v \circ \pi) = \overline{\overline{H}}_S(v).$$

5. **Expansibilité.** Soit v une bi-capacité normalisée sur N . Si $k \in N$ est un *élément nul* pour v , c.-à-d. si $v(S \cup k, T) = v(S, T) = v(S, T \cup k)$ pour tout $(S, T) \in \mathcal{Q}(N \setminus k)$, alors

$$\overline{\overline{H}}_S(v) = \overline{\overline{H}}_S(v^{N \setminus k}),$$

où $v^{N \setminus k}$ désigne la restriction de v à $\mathcal{Q}(N \setminus k)$.

6. **Décisivité.** Pour toute bi-capacité $v \in \mathcal{BC}_N^*$, nous avons

$$\overline{\overline{H}}_S(v) \geq 0.$$

De plus, $\overline{\overline{H}}_S(v) = 0$ si et seulement si, pour tout $x \in [-1, 1]^n$, il existe $\lambda \in]0, 2]$ et $i \in N$ tels que $C_v(x) = \lambda x_i$.

7. **Maximalité.** Pour toute bi-capacité $v \in \mathcal{BC}_N^*$, nous avons

$$\overline{\overline{H}}_S(v) \leq \ln n,$$

avec égalité si et seulement si v est la bi-capacité uniforme v^* sur N .

8. **Monotonie croissante vers v^* .** Soit v une bi-capacité normalisée sur N telle que $v \neq v^*$ et, pour tout $\lambda \in [0, 1]$, soit v_λ une bi-capacité normalisée sur N telle que $v_\lambda := v + \lambda(v^* - v)$. Alors, pour tout $0 \leq \lambda_1 < \lambda_2 \leq 1$, nous avons

$$\overline{\overline{H}}_S(v_{\lambda_1}) < \overline{\overline{H}}_S(v_{\lambda_2}).$$

Comme dans le cas unipolaire, en pratique il est plus approprié de calculer la version normalisée de $\overline{\overline{H}}_S(\mu)$ définie par $\overline{\overline{H}}_S^*(\mu) := \overline{\overline{H}}_S(\mu) / \ln n$. Une valeur de $\overline{\overline{H}}_S^*(\mu)$ proche de 1 indique que l'intégrale de Choquet bipolaire se comporte presque comme la moyenne arithmétique simple. À l'opposée, une valeur de $\overline{\overline{H}}_S^*(\mu)$ égale à 0 correspond à une agrégation dictatoriale car seule une évaluation partielle est utilisée dans le calcul de la valeur agrégée.

2.5 Loi et moments de l'intégrale de Choquet dans le cas uniforme

Lorsque l'intégrale de Choquet est utilisée comme modèle décisionnel (cf. Chapitre 3), son interprétation, essentiellement fondée sur les indices présentés dans les sections précédentes, peut être complétée par la connaissance de sa loi et de ses moments. Les premiers travaux dans ce domaine sont dus à Marichal (1998) qui a déterminé l'espérance et la variance de l'intégrale de Choquet dans le cas où le vecteur d'évaluations partielles est uniformément distribué sur $[0, 1]^n$. Ces résultats ont été récemment généralisés par Marichal & Kojadinovic (2006) qui ont déterminé la loi exacte ainsi que tous les moments de l'intégrale de Choquet⁵ dans le cas uniforme. La détermination de la loi dans le cas général semble extrêmement ardue. Notons que le cas gaussien a été récemment empiriquement étudié par Grabisch & Raufaste (2006).

Les résultats de Marichal & Kojadinovic (2006) présentés dans cette section peuvent être vus comme des généralisations de résultats existant sur la distribution de combinaisons linéaires de statistiques d'ordre dans le cas uniforme (Govindarajulu 1968, Matsunawa 1985, Agarwal, Dalpatadu & Singh 2002). Ils reposent sur des outils mathématiques habituellement utilisés en analyse numérique tels que les *différences divisées* et les *B-splines* (de Boor 1972, Davis 1975, Powell 1981).

2.5.1 Définitions : fonctions puissances tronquées et différences divisées

La fonction puissance tronquée $(x)_+^n$ est égale à x^n si $x > 0$ et à 0 sinon. De façon similaire, $(x)_-^n$ vaut x^n si $x < 0$ et 0 sinon.

La n -ième *différence divisée* d'une fonction g à valeurs dans \mathbb{R} est une fonction symétrique de $n + 1$ variables définie récursivement par

$$\Delta[g : a_0] := g(a_0)$$

et

$$\Delta[g : a_0, \dots, a_n] := \frac{\Delta[g : a_0, \dots, a_{n-1}] - \Delta[g : a_1, \dots, a_n]}{a_0 - a_n}$$

lorsque $a_0, \dots, a_n \in \mathbb{R}$ sont distincts. Si $a_0 = a_n$, la convention suivante est usuellement adoptée :

$$\Delta[g : a_0, \dots, a_{n-1}, a_0] := \frac{\partial}{\partial a_0} \Delta[g : a_0, \dots, a_{n-1}].$$

Dans le cas où a_0, \dots, a_n sont distincts, il peut être vérifié par récurrence que

$$\Delta[g : a_0, \dots, a_n] = \sum_{i=0}^n \frac{g(a_i)}{\prod_{\substack{j=0 \\ j \neq i}}^n (a_i - a_j)}.$$

2.5.2 Principaux résultats

Théorème 2.5.1 *Soit μ un jeu sur N et soit X un vecteur aléatoire uniformément distribué sur $[0, 1]^n$. La fonction de répartition de la variable aléatoire $C_\mu(X)$, notée F_{C_μ} , est donnée par*

$$F_{C_\mu}(y) = \frac{1}{n!} \sum_{\sigma \in \Pi_N} \Delta[(\cdot - y)_-^n : 0, \mu_n^\sigma, \dots, \mu_1^\sigma], \quad y \in \mathbb{R},$$

⁵L'extension à l'intégrale de Choquet bipolaire est immédiate.

où $\mu_i^\sigma := \mu(\{\sigma(i), \dots, \sigma(n)\})$ pour tout $i \in N$ et tout $\sigma \in \Pi_N$, avec la convention $\mu_{n+1}^\sigma := 0$. La densité de $C_\mu(X)$, notée f_{C_μ} , est donnée par

$$f_{C_\mu}(y) = \frac{1}{(n-1)!} \sum_{\sigma \in \Pi_N} \Delta[(\cdot - y)_+^{n-1} : 0, \mu_n^\sigma, \dots, \mu_1^\sigma], \quad y \in \mathbb{R}.$$

Le corollaire ci-après donne la fonction de répartition de l'intégrale de Choquet dans le cas uniforme sous la forme d'une formule fermée lorsque le jeu satisfait certaines conditions de régularité.

Corollaire 2.5.1 *Soit μ un jeu sur N tel que, pour tout $\sigma \in \Pi_N$, les nombres $(\mu_i^\sigma)_{i \in N}$ soient tous distincts, et soit X un vecteur aléatoire uniformément distribué sur $[0, 1]^n$. La fonction de répartition de la variable aléatoire $C_\mu(X)$ est donnée par*

$$F_{C_\mu}(y) = \frac{1}{n!} \sum_{\sigma \in \Pi_N} \sum_{i=1}^{n+1} \frac{(\mu_i^\sigma - y)_+^n}{\prod_{\substack{k=1 \\ k \neq i}}^{n+1} (\mu_i^\sigma - \mu_k^\sigma)}, \quad y \in \mathbb{R}.$$

Les propositions suivantes donnent la fonction génératrice des moments ainsi que les moments de l'intégrale de Choquet dans le cas uniforme.

Proposition 2.5.1 *Soit μ un jeu sur N et soit X un vecteur aléatoire uniformément distribué sur $[0, 1]^n$. La fonction génératrice des moments de $C_\mu(X)$ est donnée par*

$$\mathcal{M}_{C_\mu}(t) = \frac{1}{t^n} \sum_{\sigma \in \Pi_N} \Delta[e^{t(\cdot)} : 0, \mu_n^\sigma, \dots, \mu_1^\sigma], \quad t \in \mathbb{R}.$$

Proposition 2.5.2 *Soit μ un jeu sur N et soit X un vecteur aléatoire uniformément distribué sur $[0, 1]^n$. Pour tout entier $r > 0$, le moment d'ordre r de $C_\mu(X)$ est donné par*

$$E[C_\mu(X)^r] = \frac{r!}{(n+r)!} \sum_{\sigma \in \Pi_N} \Delta[(\cdot)^{n+r} : 0, \mu_n^\sigma, \dots, \mu_1^\sigma],$$

et le moment centré d'ordre r par

$$E[(C_\mu(X) - E[C_\mu(X)])^r] = \frac{r!}{(n+r)!} \sum_{\sigma \in \Pi_N} \Delta[(\cdot - E[Y])^{n+r} : 0, \mu_n^\sigma, \dots, \mu_1^\sigma].$$

Le calcul pratique de différences divisées de fonctions puissances n'est pas problématique (voir p. ex. Ali 1973). En revanche, le calcul de différences divisées de fonctions puissances tronquées est plus complexe.

2.5.3 Calcul de différences divisées de fonctions puissances tronquées

Le calcul des quantités $F_{C_\mu}(y)$ et $f_{C_\mu}(y)$ requiert le calcul de différences divisées de fonctions puissances tronquées, aussi connues sous le nom de *B-splines* (de Boor 1972, Davis 1975, Powell 1981). Tous les algorithmes existants permettant le calcul de *B-splines* sont fondés sur la même relation de récurrence.

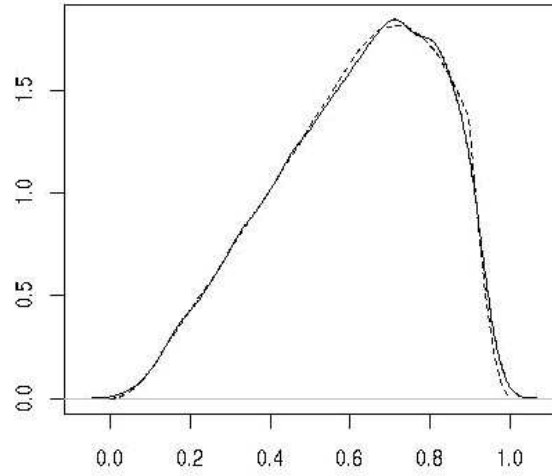


FIG. 2.1 – En pointillés, densité exacte de l'intégrale de Choquet par rapport à μ . En trait continu, densité estimée par la méthode des noyaux à partir de 10 000 réalisations obtenues par génération pseudo-aléatoire.

Soit $y \in \mathbb{R}$ et soient $a_0, \dots, a_n \in \mathbb{R}$ non nécessairement tous distincts. Notons b_1, \dots, b_r les a_i tels que $a_i < y$ et notons c_1, \dots, c_s les a_i tels que $a_i \geq y$. Nous avons ainsi nécessairement $r + s = n + 1$. Il peut alors être vérifié que

$$\alpha_{k,l} = \frac{(c_l - y)\alpha_{k-1,l} + (y - b_k)\alpha_{k,l-1}}{c_l - b_k}, \quad (2.9)$$

où

$$\alpha_{k,l} := \Delta \left[(\cdot - y)_+^{k+l-2} : b_1, \dots, b_k, c_1, \dots, c_l \right], \quad k + l \geq 2.$$

L'Eq. (2.9) est une forme équivalente de la formule de récurrence bien connue à la base de l'*algorithme de de Boor (1972)* pour le calcul des B -splines. Avec les conditions initiales $\alpha_{1,1} = (c_1 - b_1)^{-1}$ et $\alpha_{0,l} = \alpha_{k,0} = 0$ pour tous $l, k \geq 2$, la récurrence précédente permet le calcul de $\alpha_{r,s} = \Delta[(\cdot - y)_+^{n-1} : a_0, \dots, a_n]$ en $O(n^2)$ opérations. De façon similaire, il peut être vérifié que la suite

$$\beta_{k,l} := \Delta \left[(\cdot - y)_-^{k+l-1} : b_1, \dots, b_k, c_1, \dots, c_l \right], \quad k + l \geq 1,$$

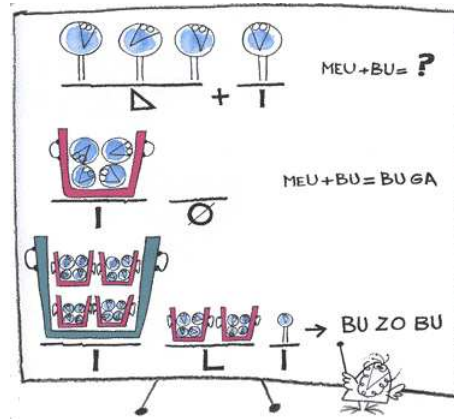
satisfait aussi la relation de récurrence donnée dans l'Eq. (2.9) (Ali 1973, Varsi 1973). Avec les conditions initiales $\beta_{0,l} = 0$ pour tout $l \geq 1$ et $\beta_{k,0} = 1$ pour tout $k \geq 1$, la formule de récurrence ci-dessus permet le calcul de $\beta_{r,s} = \Delta[(\cdot - y)_-^n : a_0, \dots, a_n]$ en $O(n^2)$ opérations.

2.5.4 Exemple

Le calcul de la fonction de répartition et de la densité de l'intégrale de Choquet dans le cas uniforme a été implémenté dans le *package* Kappalab (Grabisch, Kojadinovic & Meyer 2005) pour GNU R (R Development Core Team 2005) qui sera présenté de façon plus détaillée dans le chapitre suivant (voir aussi la Section A).

À titre d'exemple, considérons la capacité normalisée μ sur $\{1, 2, 3\}$ dont les coefficients sont $\mu(1) = 0.1$, $\mu(2) = 0.6$, $\mu(3) = \mu(12) = \mu(13) = \mu(23) = 0.9$. La densité de l'intégrale de Choquet par rapport à μ dans le cas uniforme est représentée dans la Figure 2.1 en pointillés. Le trait continu correspond à la densité estimée par la méthode des noyaux à partir de 10 000 réalisations obtenues par génération pseudo-aléatoire.

Identification de modèles décisionnels fondés sur l'intégrale de Choquet



Jacques Rouxel, *Les Shadoks*.

Sommaire

3.1	Introduction	46
3.2	Théorie de l'utilité multiattribut fondée sur l'intégrale de Choquet	47
3.2.1	Contexte théorique	47
3.2.2	Formulation du problème d'identification de capacités	48
3.3	Principales méthodes d'identification de capacités	50
3.3.1	Approches fondées sur les moindres carrés	50
3.3.2	Approche fondée sur la programmation linéaire	51
3.3.3	Méthodes du minimum de variance et du minimum de distance	52
3.3.4	Moindres carrés généralisés	53
3.3.5	Implémentation des méthodes d'identification	54
3.4	Exemples d'utilisation des méthodes d'identification	55
3.4.1	Description du problème	55
3.4.2	Approches fondées sur les moindres carrés	56
3.4.3	Approches linéaire, du minimum de variance et du minimum de distance	58

3.4.4	Contraintes supplémentaires sur la valeur de Shapley	60
3.4.5	Contraintes supplémentaires sur les indices d'interaction	62
3.4.6	Solution plus simple mais approchée	64

3.1 Introduction

Les travaux fondateurs de Schmeidler (1989) semblent être à l'origine des premières applications de l'intégrale de Choquet. Dans le domaine de la décision dans le risque et l'incertain, Schmeidler a proposé d'utiliser l'intégrale de Choquet afin de généraliser le *modèle d'utilité espérée* classique (voir aussi Tversky & Kahneman 1992, Chateauneuf 1994, Narukawa & Murofushi 2004). Les études menées sur les intégrales non additives, et leurs caractérisations axiomatiques en tant qu'opérateurs d'agrégation, ont conduit plus récemment à l'utilisation de l'intégrale de Choquet en aide multicritère à la décision (voir p. ex. Grabisch 1992, Grabisch 1995a, Marichal 2000a, Marichal 2000b, Grabisch & Labreuche 2004) et en discrimination (Grabisch & Nicolas 1994, Grabisch 1996).

Quel que soit le domaine d'application, l'utilisation de l'intégrale de Choquet dans un modèle décisionnel requiert au préalable l'*identification* d'une capacité. L'objectif de ce chapitre, fortement fondé sur (Grabisch, Kojadinovic & Meyer 2006), est de présenter un état de l'art des principales méthodes d'identification existant dans la littérature. Afin d'offrir une présentation unifiée des approches existantes, nous avons choisi de nous placer dans le contexte de l'*approche cardinale* de l'aide multicritère à la décision connue sous le nom de *théorie de l'utilité multiattribut* (Keeney & Raiffa 1976, Vincke 1992).

La plupart des méthodes d'identification de capacités proposées dans la littérature peuvent se mettre sous la forme de problèmes d'optimisation. Elles diffèrent selon leur fonction objectif et le type d'information préférentielle qu'elles nécessitent en entrée. Pour chacune des méthodes présentées, nous donnerons ses principaux avantages et inconvénients. La dernière section de ce chapitre sera consacrée à des exemples d'application de ces méthodes et à leur mise en œuvre dans le *package* Kappalab (Grabisch, Kojadinovic & Meyer 2005). Kappalab (pour *laboratory for capacities* en anglais) est un ajout pour le système statistique libre GNU R (R Development Core Team 2005) qui peut être vu comme un Matlab pour les statistiques et l'analyse de données. En plus des routines d'identification de capacités mentionnées, Kappalab contient des fonctions pour la manipulation de jeux, de capacités et d'intégrales non additives qui peuvent être également utilisées dans le contexte de la théorie des jeux coopératifs (cf. Section A).

La première section de ce chapitre est consacrée à une présentation de la théorie de l'utilité multiattribut fondée sur l'intégrale de Choquet. Dans la deuxième section, nous décrirons succinctement les principales méthodes d'identification de capacités proposées dans la littérature ; nous ferons en particulier le lien avec la Section 2.4.3 du chapitre précédent en présentant les méthodes du *minimum de variance* (Kojadinovic 2006a) et du *minimum de distance* (Kojadinovic 2006b). Des exemples d'utilisation des méthodes décrites seront présentés dans la dernière section et leur mise en œuvre pratique dans le *package* Kappalab sera détaillée.

3.2 Théorie de l'utilité multiattribut fondée sur l'intégrale de Choquet

3.2.1 Contexte théorique

Soit $X := X_1 \times \cdots \times X_n$, $n \geq 2$, un ensemble d'objets d'intérêt décrits par un ensemble $N := \{1, \dots, n\}$ d'attributs⁶. L'objectif de la théorie de l'utilité multiattribut (Keeney & Raiffa 1976) est de modéliser *numériquement* les préférences⁷ d'un décideur. Ces préférences prennent mathématiquement la forme d'une relation binaire \succeq qu'il s'agit de représenter par le biais d'une *fonction d'utilité globale* $U : X \rightarrow \mathbb{R}$ telle que

$$x \succeq y \iff U(x) \geq U(y), \quad \forall x, y \in X.$$

La relation de préférence \succeq est usuellement supposée complète et transitive (voir p. ex. Bouyssou & Pirlot 2006, pour le cas de préférences non transitives ou non complètes). Le modèle le plus fréquemment utilisé pour la fonction d'utilité globale est le modèle d'*utilité additive* (voir p. ex. Bouyssou & Pirlot 2005, Gonzales 1996). Dans ce qui suit, nous considérons le modèle plus général, dit *transitif décomposable*, de Krantz, Luce, Suppes & Tversky (1971) (voir aussi Bouyssou & Pirlot 2004), dans lequel U est définie par

$$U(x) := F(u_1(x_1), \dots, u_n(x_n)), \quad \forall x = (x_1, \dots, x_n) \in X, \quad (3.1)$$

où les fonctions $u_i : X_i \rightarrow \mathbb{R}$ sont appelées les *fonctions d'utilité* et $F : \mathbb{R}^n \rightarrow \mathbb{R}$, non-décroissante en chacun de ses arguments, est parfois appelée la *fonction d'agrégation*.

Les fonctions d'utilité peuvent être vues comme permettant de traduire les valeurs d'un attribut en degrés de satisfaction : pour tout $x \in X$, la quantité $u_i(x_i)$ peut être interprétée comme une mesure de la satisfaction du décideur relativement à la valeur x_i de l'attribut i . Dans la suite, le terme *critère* sera utilisé pour désigner l'association d'un attribut $i \in N$ avec la fonction d'utilité correspondante u_i . Une condition nécessaire à l'existence du modèle décomposable ci-dessus est que la relation de préférence \succeq soit un préordre *faiblement séparable* (voir p. ex. Bouyssou & Pirlot 2005).

La forme exacte de la fonction d'utilité globale U dépend des hypothèses sur lesquelles le problème d'aide multicritère à la décision est fondé. Lorsque que l'*indépendance préférentielle* entre critères peut être supposée (voir p. ex. Vincke 1992, Gonzales 1996, Bouyssou & Pirlot 2006), il est fréquent de considérer que la fonction F est additive et prend la forme d'une somme pondérée. Le modèle décomposable donné par l'Eq. (3.1) est alors équivalent au modèle additif classique. L'hypothèse d'indépendance préférentielle est cependant rarement vérifiée en pratique. Afin de permettre la prise en compte de phénomènes d'interaction entre critères, il a été récemment suggéré de remplacer la somme pondérée par une intégrale de Choquet (voir p. ex. Grabisch 1992, Marichal 2000a, Grabisch & Labreuche 2004).

L'utilisation de l'intégrale de Choquet en tant que fonction d'utilité globale dans l'Eq. (3.1) nécessite, comme nous l'avons vu dans la Section 1.3.1, la possibilité de comparer le degré d'utilité d'une alternative selon différents critères. En d'autres termes, il est nécessaire que les fonctions

⁶Tous les éléments de X ne correspondent pas nécessairement à des objets réels. Les objets correspondant à des combinaisons impossibles ou improbables de valeurs d'attributs sont usuellement dits *fictifs*.

⁷La théorie de l'utilité multiattribut est une application du *mesurage conjoint* (voir p. ex. Wakker 1989, Bouyssou & Pirlot 2006) à l'aide multicritère à la décision.

d'utilité soient *commensurables*, c.-à-d. $u_i(x_i) = u_j(x_j)$ si et seulement si, pour le décideur, l'alternative x est satisfaite avec la même intensité selon les critères i et j (voir p. ex. Grabisch, Labreuche & Vansnick 2003, Grabisch 2005b, pour une discussion plus approfondie sur la notion de commensurabilité).

Dans le contexte considéré, Labreuche & Grabisch (2003) ont proposé une construction rigoureuse des fonctions d'utilité fondée sur la méthode MACBETH (Bana e Costa & Vansnick 1999, Bana e Costa, Corte & Vansnick 2005). Nous renvoyons le lecteur intéressé à (Grabisch & Labreuche 2004, Grabisch 2005b) car cet aspect ne sera pas abordé dans ce chapitre.

Dans la suite, étant donnée une relation binaire \succeq , nous noterons \succ sa partie asymétrique et \sim sa partie symétrique. De plus, pour tout $x \in X$, le vecteur des utilités partielles de x sera noté $u(x)$, c.-à-d. $u(x) := (u_1(x_1), \dots, u_n(x_n))$.

3.2.2 Formulation du problème d'identification de capacités

Une fois des fonctions d'utilité commensurables déterminées, l'étape suivante consiste à identifier une capacité, si elle existe, telle que l'intégrale de Choquet par rapport à cette capacité représente numériquement les préférences du décideur. Dans le cadre du traitement pratique d'un problème d'aide multicritère à la décision, le décideur raisonne généralement sur un sous-ensemble \mathcal{O} , usuellement de faible cardinal, de l'ensemble X des objets d'intérêts. Les *préférences initiales* du décideur, à partir desquelles la capacité doit être déterminée, peuvent prendre la forme :

- d'un préordre partiel $\succeq_{\mathcal{O}}$ sur \mathcal{O} (rangement des objets disponibles) ;
- d'un préordre partiel \succeq_N sur N (rangement des critères) ;
- d'intuitions quantitatives sur l'importance de certains critères ;
- d'un préordre partiel \succeq_P sur l'ensemble des paires de critères (rangement des *interactions entre critères*) ;
- d'intuitions sur le type et l'intensité de l'interaction entre certains critères ;
- du caractère de *veto* ou de *faveur* de certains critères (Grabisch 1997a, Marichal 2000b, Marichal 2004b) ;
- d'une partition *inter-additive* de N , c.-à-d. grosso modo une partition des critères en classes indépendantes deux à deux (Fujimoto & Murofushi 2000) ;
- etc.

Dans le contexte de la théorie de l'utilité multiattribut fondée sur l'intégrale de Choquet, il semble naturel de traduire certaines des préférences initiales exprimées ci-dessus comme suit :

- $x \succ_{\mathcal{O}} x'$ peut être traduit par $C_{\mu}(u(x)) - C_{\mu}(u(x')) \geq \delta_C$;
- $x \sim_{\mathcal{O}} x'$ peut être traduit par $-\delta_C \leq C_{\mu}(u(x)) - C_{\mu}(u(x')) \leq \delta_C$;
- $i \succ_N j$ peut être traduit par $\phi_{\mu}(i) - \phi_{\mu}(j) \geq \delta_{Sh}$;
- $i \sim_N j$ peut être traduit par $-\delta_{Sh} \leq \phi_{\mu}(i) - \phi_{\mu}(j) \leq \delta_{Sh}$;
- $ij \succ_P kl$ peut être traduit par $I_{\mu}(ij) - I_{\mu}(kl) \geq \delta_I$;
- $ij \sim_P kl$ peut être traduit par $-\delta_I \leq I_{\mu}(ij) - I_{\mu}(kl) \leq \delta_I$;

où μ est la capacité à déterminer et où δ_C , δ_{Sh} et δ_I sont des seuils d'indifférence fixés par le décideur (voir p. ex. Marichal & Roubens 2000a, Kojadinovic 2006a).

Bien que cela soit plus discutable, les informations restantes de nature plus quantitative pourraient être traduites comme suit :

- une intuition quantitative sur l'importance relative d'un critère i pourrait être traduite par $a \leq \phi_{\mu}(i) \leq b$, où $a, b \in [0, 1]$, $a \leq b$, doivent être fixés par le décideur ;
- une intuition sur le type et l'intensité de l'interaction entre deux critères i et j pourrait

être traduite par $a \leq I_\mu(ij) \leq b$, où $0 \leq a \leq b \leq 1$ dans le cas d'une complémentarité entre critères et où $-1 \leq a \leq b \leq 0$ dans le cas de critères substitutifs.

Notons qu'en pratique une contrainte de la forme $I_\mu(ij) - I_\mu(kl) \geq \delta_I$ est généralement accompagnée soit par la contrainte $0 \geq I_\mu(ij)$ ou par la contrainte $I_\mu(kl) \geq 0$.

Enfin, le comportement en tant que veto (resp. faveur) d'un critère i peut être directement traduit par $\mu(T) = 0$ pour tout $T \subseteq N$ tel que $T \not\ni k$ (resp. $\mu(T) = 1$ pour tout $T \subseteq N$ tel que $T \ni k$) (Marichal 2000b, Prop. 3 et 4).

La plupart des méthodes d'identification proposées dans la littérature peuvent alors se mettre sous la forme du problème d'optimisation suivant :

$$\begin{array}{ll} \min \text{ ou } \max \mathcal{F}(\dots) & \\ \text{sous les contraintes} & \left\{ \begin{array}{l} \nu(S \cup i) - \nu(S) \geq 0, \forall i \in N, \forall S \subseteq N \setminus i, \\ \nu(N) = 1, \\ C_\nu(u(x)) - C_\nu(u'(x)) \geq \delta_C, \\ \vdots \\ \phi_\nu(i) - \phi_\nu(j) \geq \delta_{Sh}, \\ \vdots \\ I_\nu(ij) - I_\nu(kl) \geq \delta_I, \\ I_\nu(kl) \geq 0 \\ \vdots \end{array} \right. \end{array}$$

où ν est un jeu sur N et \mathcal{F} est une fonction objectif qui diffère selon la méthode d'identification choisie.

Une solution au problème précédent est une capacité déterminée par $2^n - 1$ coefficients. Nous voyons ainsi que le nombre de variables impliquées dans le problème croît exponentiellement avec n . Il en est de même du temps de résolution. En pratique, autant pour des raisons relatives au coût de calcul que par souci de simplicité, il est préférable de restreindre l'ensemble des solutions possibles aux capacités k -additives, k étant fixé dans $\{1, \dots, n\}$, typiquement $k = 2$ ou 3 . Il semble alors naturel de récrire le problème d'optimisation ci-dessus en fonction de la transformée de Möbius d'un jeu k -additif en utilisant la Proposition 1.2.1 et les Eqs. (1.5), (2.1) et (2.3), ce qui a pour effet de faire passer le nombre de variables de $2^n - 1$ à $\sum_{l=1}^k \binom{n}{l}$. Nous obtenons ainsi

$$\begin{array}{ll} \min \text{ ou } \max \mathcal{F}(\dots) & \\ \text{sous les contraintes} & \left\{ \begin{array}{l} \sum_{\substack{T \subseteq S \\ t \leq k-1}} m_\nu(T \cup i) \geq 0, \forall i \in N, \forall S \subseteq N \setminus i, \\ \sum_{\substack{T \subseteq N \\ 0 < t \leq k}} m_\nu(T) = 1, \\ C_{m_\nu}(u(x)) - C_{m_\nu}(u'(x)) \geq \delta_C, \\ \vdots \\ \phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq \delta_{Sh}, \\ \vdots \\ I_{m_\nu}(ij) - I_{m_\nu}(kl) \geq \delta_I, \\ I_{m_\nu}(kl) \geq 0 \\ \vdots \end{array} \right. \end{array}$$

où m_ν désigne la représentation de Möbius d'un jeu k -additif ν sur N , et où C_{m_ν} , ϕ_{m_ν} et I_{m_ν} désignent les expressions de l'intégrale de Choquet, de la valeur de Shapley et des indices d'interaction respectivement en fonction de la représentation de Möbius de ν données dans les Eqs. (1.5), (2.1) et (2.3).

Le problème d'optimisation précédent n'aura pas de solution si les contraintes sont inconsistantes. Cela peut être du soit au fait que les informations préférentielles fournies par le décideur sont contradictoires, soit parce que le nombre de paramètres du modèle, c.-à-d. le nombre de coefficients de la transformée de Möbius, est trop faible pour que toutes les contraintes puissent être satisfaites. Dans ce dernier cas, afin d'augmenter le nombre de variables, la pratique consiste usuellement à incrémenter l'ordre de k -additivité. Il est néanmoins possible que les contraintes imposées par le décideur ne puissent pas être satisfaites même en utilisant une capacité n -additive. La conclusion qui s'impose alors est que l'intégrale de Choquet n'est pas un opérateur d'agrégation suffisamment flexible pour modéliser les préférences initiales du décideur (voir p. ex. Wakker 1989). Le recours à l'intégrale de Choquet par rapport à une bi-capacité peut être une solution si l'échelle d'utilité s'avère en réalité bipolaire.

3.3 Principales méthodes d'identification de capacités

Après une présentation des méthodes d'identification fondées sur les moindres carrés pouvant être vues comme des généralisations de la régression linéaire multiple, nous décrirons une approche simple, reposant sur la programmation linéaire, deux approches quadratiques, privilégiant les solutions “moyennes”, et une méthode plus souple, permettant l'obtention de solutions ne vérifiant pas nécessairement toutes les contraintes imposées par le décideur.

3.3.1 Approches fondées sur les moindres carrés

Mori & Murofushi (1989) sont probablement les premiers à avoir proposé une méthode d'identification de capacités. L'approche qu'ils ont suggérée peut être vue comme une généralisation de la régression linéaire multiple. Sa mise en œuvre requiert en effet additionnellement la donnée d'utilités globales “objectifs” pour les objets disponibles. Ces utilités, fournies par le décideur, seront notées $(y(x))_{x \in \mathcal{O}}$ dans la suite. La fonction objectif du problème d'optimisation est alors définie par

$$\mathcal{F}_{MC}(m_\nu) := \sum_{x \in \mathcal{O}} [C_{m_\nu}(u(x)) - y(x)]^2,$$

Le but de la méthode est donc de déterminer la capacité ν qui minimise la distance quadratique entre les utilités globales $(C_{m_\nu}(u(x)))_{x \in \mathcal{O}}$ calculées par le biais de l'intégrale de Choquet par rapport à ν et les utilités globales $(y(x))_{x \in \mathcal{O}}$ souhaitées par le décideur. Le problème d'optimisation prend ainsi la forme d'un programme quadratique, non nécessairement strictement convexe (Grabisch, Nguyen & Walker 1995, Miranda & Grabisch 1999), ce qui implique que la solution, si elle existe, n'est pas nécessairement unique. Afin d'éviter l'utilisation de solveurs quadratiques, des approches heuristiques ont été proposées par Mori & Murofushi (1989), Ishii & Sugeno (1996) et Grabisch (1995b). Nous donnons ci-après les grandes lignes de cette dernière méthode, appelée *moindres carrées heuristiques*.

L'approche proposée par Grabisch (1995b) est une méthode du gradient ayant pour point de départ une capacité normalisée μ définie par le décideur que nous appellerons la *capacité*

initiale. Cette capacité, typiquement additive, représente l'idée que le décideur se fait a priori de la fonction d'agrégation F dans l'Eq. (3.1). En l'absence de prérequis clairs sur la fonction d'agrégation, un choix très naturel pour μ est la capacité uniforme μ^* car l'intégrale de Choquet par rapport à μ^* n'est rien d'autre que la moyenne arithmétique simple. Une fois la capacité initiale μ choisie, pour un objet $x \in \mathcal{O}$, le gradient ne modifie que les coefficients de μ impliqués dans le calcul de $C_\mu(u(x))$, dans les limites imposées par la monotonie de μ . Une fois tous les objets disponibles utilisés, les coefficients non modifiés de μ sont rapprochés de la valeur moyenne des coefficients voisins relativement au treillis $(\mathcal{P}(N), \subseteq)$ (cf. Section 1.2.6). Les étapes précédentes forment une itération et sont répétées jusqu'à ce qu'un critère d'arrêt soit satisfait. L'avantage sur la méthode quadratique optimale est que seul le vecteur des coefficients de μ doit être stocké, alors que, dans cette dernière, une matrice carrée de dimensions similaires doit être mise en mémoire. De plus, comme nous le verrons dans l'exemple de la Section 3.4.2, cette approche heuristique conduit généralement à des solutions moins extrêmes que l'approche optimale. Néanmoins, contrairement à la méthode quadratique optimale, il n'est pas possible de restreindre l'ensemble des solutions aux capacités k -additives, $1 \leq k < n$.

Dans le contexte de la théorie de l'utilité multiattribut fondée sur l'intégrale de Choquet, l'inconvénient principal des approches présentées ci-dessus est qu'elles requièrent la connaissance supplémentaire des utilités globales souhaitées $(y(x))_{x \in \mathcal{O}}$, lesquelles ne peuvent pas toujours être fournies par le décideur.

3.3.2 Approche fondée sur la programmation linéaire

Marichal & Roubens (2000a) ont proposé une méthode d'identification simple, fondée sur la programmation linéaire. L'approche proposée peut être formulée comme suit :

$$\begin{aligned} \max \mathcal{F}_{PL}(\varepsilon) &:= \varepsilon \\ \text{sous les contraintes} &\left\{ \begin{array}{l} \sum_{\substack{T \subseteq S \\ t \leq k-1}} m_\nu(T \cup i) \geq 0, \forall i \in N, \forall S \subseteq N \setminus i, \\ \sum_{\substack{T \subseteq N \\ 0 < t \leq k}} m_\nu(T) = 1, \\ C_{m_\nu}(u(x)) - C_{m_\nu}(u(x')) \geq \delta_C + \varepsilon, \\ \vdots \\ \phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq \delta_{Sh}, \\ \vdots \end{array} \right. \end{aligned}$$

Comme nous pouvons le voir, il s'agit grosso modo de maximiser la différence minimale entre les utilités globales d'objets qui ont été rangés par le décideur par le biais du préordre partiel $\succeq_{\mathcal{O}}$. En effet, si le décideur estime que $x \succ_{\mathcal{O}} x'$, c'est que x est significativement meilleur que x' . Il peut alors sembler naturel de souhaiter que leurs utilités globales reflètent cette différence de la façon la plus significative possible.

L'avantage principal de cette approche est sa simplicité. Néanmoins, comme dans le cas des approches proposées dans la section précédente, cette méthode d'identification ne conduit pas nécessairement à une solution unique, si elle existe. De plus, comme nous le verrons dans la Section 3.4, la solution obtenue peut parfois être considérée comme trop extrême car elle correspond à une capacité qui maximise la différence entre objets rangés par le décideur.

Notons que cette approche a été étendue afin de permettre la résolution de problèmes multicritère de tri (Marichal, Meyer & Roubens 2005).

3.3.3 Méthodes du minimum de variance et du minimum de distance

L'idée de la méthode du *minimum de variance* (Kojadinovic 2006a) est de favoriser la capacité la “moins spécifique” compatible avec les préférences initiales du décideur, si elle existe. La fonction objectif est prise égale à la *variance* de la capacité (cf. Eq. (2.7)), c.-à-d.

$$\mathcal{F}_{MV}(m_\nu) := \frac{1}{n} \sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) \left(\sum_{T \subseteq S} m_\mu(T \cup i) - \frac{1}{n} \right)^2.$$

Comme nous l'avons vu dans la Section 2.4.3, minimiser la variance est équivalent à maximiser l'entropie étendue de Havrda & Charvat d'ordre 2. Cette méthode peut donc être regardée de façon équivalente comme une approche du *maximum d'entropie*⁸.

Le problème d'optimisation prend la forme du programme quadratique strictement convexe suivant :

$$\begin{aligned} & \min \mathcal{F}_{MV}(m_\nu) \\ & \text{sous les contraintes} \left\{ \begin{array}{l} \sum_{\substack{T \subseteq S \\ t \leq k-1}} m_\nu(T \cup i) \geq 0, \forall i \in N, \forall S \subseteq N \setminus i, \\ \sum_{\substack{T \subseteq N \\ 0 < i \leq k}} m_\nu(T) = 1, \\ C_{m_\nu}(u(x)) - C_{m_\nu}(u(x')) \geq \delta_C, \\ \vdots \\ \phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq \delta_{Sh}, \\ \vdots \end{array} \right. \end{aligned}$$

Comme évoqué dans (Kojadinovic, Marichal & Roubens 2005, Kojadinovic 2006a) et dans la Section 2.4.4, l'intégrale de Choquet par rapport à la capacité de variance minimale compatible avec les préférences initiales du décideur, si elle existe, est celle qui va exploiter le plus en moyenne ses arguments.

Un des avantages de cette approche est qu'elle conduit à une solution unique, si elle existe, à cause de la stricte convexité de la fonction objectif. De plus, dans le cas où les préférences initiales du décideur sont “pauvres”, c.-à-d. se traduisent par un faible nombre de contraintes, cette solution unique n'aura pas des caractéristiques trop extrêmes comme des interactions très fortes en valeur absolue entre certains critères ou des critères prenant une importance démesurée dans la décision.

Une généralisation de cette approche, proposée par Kojadinovic (2006b), consiste à trouver, si elle existe, la capacité la plus proche d'une capacité définie par le décideur et compatible avec ses préférences initiales. Cette *capacité initiale* est typiquement additive et représente l'idée que le décideur se fait a priori de la fonction d'agrégation. En l'absence de prérequis clairs de la part du décideur, comme déjà évoqué dans la Section 3.3.1, un choix très naturel pour μ est la capacité uniforme μ^* . Afin de pouvoir implémenter pratiquement un tel principe, dans

⁸Il faut néanmoins rappeler, comme nous l'avons évoqué dans la Section 2.4, que la notion d'*entropie* n'est pas ici liée à celle d'*incertitude*.

(Kojadinovic 2006b), trois distances quadratiques ont été étudiées. Dans la suite, nous nous limitons à la distance suivante définie, pour toute paire de jeux μ, ν sur N par

$$d^2(m_\mu, m_\nu) := \int_{[0,1]^n} [C_{m_\nu}(x) - C_{m_\mu}(x)]^2 dx. \quad (3.2)$$

Cette distance quadratique, étudiée de façon approfondie par Marichal (1998, Chap. 7) dans le contexte de l'extension de fonctions pseudo-bouliennes, peut être interprétée comme l'espérance par rapport à la distribution uniforme sur $[0, 1]^n$ de la distance quadratique entre utilités globales calculées par C_{m_μ} et utilités globales calculées par C_{m_ν} .

En l'absence de prérequis clairs sur la fonction d'agrégation, une fonction objectif naturelle pour la méthode du minimum de distance est ainsi donnée par

$$\mathcal{F}_{MD}(m_\nu) := \int_{[0,1]^n} \left[C_{m_\nu}(x) - \frac{1}{n} \sum_{i=1}^n x_i \right]^2 dx.$$

Le problème d'optimisation résultant est encore une fois un programme quadratique strictement convexe.

3.3.4 Moindres carrés généralisés

Meyer & Roubens (2005) ont proposé une généralisation des méthodes d'identification décrites dans la Section 3.3.1 plus adaptée à la théorie de l'utilité multiattribut. L'information minimale qui doit être fournie pour la mise en œuvre de l'approche proposée est un préordre total sur l'ensemble des objets disponibles et non plus leurs utilités souhaitées. La fonction objectif, faisant intervenir plus de variables que dans le cas des moindres carrés présentés dans la Section 3.3.1, est définie par

$$\mathcal{F}_{MCG}(m_\nu, y) := \sum_{x \in \mathcal{O}} [C_{m_\nu}(u(x)) - y(x)]^2,$$

où $y = (y(x))_{x \in \mathcal{O}}$ sont des variables additionnelles du programme quadratique représentant des évaluations globales inconnues des objets disponibles devant vérifier le préordre total imposé par le décideur.

Le problème d'optimisation prend la forme du programme quadratique convexe suivant :

$$\begin{array}{ll} \min \mathcal{F}_{MCG}(m_\nu, y) & \\ \text{sous les contraintes} & \left\{ \begin{array}{l} \sum_{\substack{T \subseteq S \\ t \leq k-1}} m_\nu(T \cup i) \geq 0, \forall i \in N, \forall S \subseteq N \setminus i, \\ \sum_{\substack{T \subseteq N \\ 0 < t \leq k}} m_\nu(T) = 1, \\ y(x) - y(x') \geq \delta_y, \\ \vdots \\ \phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq \delta_{Sh}, \\ \vdots \end{array} \right. \end{array}$$

où δ_y est un seuil qui peut être interprété comme la différence minimale entre les utilités globales de deux objets qui sont considérés comme significativement différents par le décideur. Une

solution au problème est composée de la représentation de Möbius m_ν de la capacité recherchée et des évaluations globales $y = (y(x))_{x \in \mathcal{O}}$ servant à traduire le préordre total du décideur.

Donnons tout d'abord une interprétation intuitive de cette approche. Comme nous l'avons évoqué dans le dernier paragraphe de la Section 3.2.2, même si les contraintes imposées par le décideur ne sont pas contradictoires, il n'est pas impossible que le nombre de paramètres du modèle (liés à l'ordre de k -additivité choisi) soit trop faible pour que les contraintes puissent être satisfaites. Une première solution consiste à incrémenter k , si possible. Une seconde solution consiste à relâcher les contraintes en traduisant le préordre sur les objets disponibles par le biais des évaluations globales inconnues $(y(x))_{x \in \mathcal{O}}$. Dans ce cas, la transformée de Möbius m_ν est moins contrainte et des solutions peuvent exister.

Le rôle de la fonction objectif est de minimiser la distance quadratique entre la représentation numérique $(y(x))_{x \in \mathcal{O}}$ du préordre imposé par le décideur et les utilités globales calculées par l'intégrale de Choquet. Si la fonction objectif est nulle, alors, pour tout objet x , son évaluation globale $y(x)$ coïncide avec l'utilité agrégée $C_{m_\nu}(u(x))$. Dans ce cas, le préordre induit par les utilités agrégées correspond au préordre imposé par le décideur et le seuil δ_y est respecté. Il peut alors exister une unique solution ou une infinité de solutions au problème. Dans le second cas, les caractéristiques de la solution qui sera choisie par le solveur sont difficiles à prévoir.

Si la fonction objectif est strictement positive, alors les utilités agrégées $(C_{m_\nu}(u(x)))_{x \in \mathcal{O}}$ ne coïncident pas exactement avec les évaluations globales inconnues $(y(x))_{x \in \mathcal{O}}$ représentant numériquement le préordre imposé par le décideur. Deux cas de figure se présentent alors : soit le préordre induit par les utilités agrégées $(C_{m_\nu}(u(x)))_{x \in \mathcal{O}}$ correspond à $\succeq_{\mathcal{O}}$ mais $x \succ_{\mathcal{O}} x'$ n'implique pas nécessairement $C_{m_\nu}(u(x)) \geq C_{m_\nu}(u(x')) + \delta_y$ (cf. Section 3.4.6), ou le préordre induit par $(C_{m_\nu}(u(x)))_{x \in \mathcal{O}}$ ne correspond pas à $\succeq_{\mathcal{O}}$. Dans le premier cas, la solution ne respecte pas le seuil minimal δ_y fixé par le décideur. Dans le second cas, le préordre est violé "en moyenne" ce qui peut conduire à une solution très peu satisfaisante.

L'avantage de cette approche est qu'elle peut fournir une solution lorsque le préordre sur les objets disponibles ne peut pas être modélisé par une intégrale de Choquet, lorsque le seuil δ_y est trop élevé, ou lorsque certaines des contraintes sur les critères ne sont pas compatibles avec une représentation du préordre du décideur par une intégrale de Choquet. Néanmoins, comme nous l'avons déjà expliqué, cette approche doit être utilisée avec précautions si la fonction objectif est nulle, car alors, cela revient à laisser le solveur quadratique choisir une solution dont les caractéristiques sont difficiles à prévoir.

3.3.5 Implémentation des méthodes d'identification

Les méthodes d'identification présentées dans les sections précédentes ont été implémentées dans le *package* Kappalab (Grabisch, Kojadinovic & Meyer 2005) pour la plateforme statistique GNU R (R Development Core Team 2005). Le *package* est distribué sous la licence libre française CeCILL de CEA, du CNRS et de l'INRIA (<http://www.cecill.info>) et peut être téléchargé sur le site du *Comprehensive R Archive Network* (<http://cran.r-project.org>) ou à partir du site qui lui est dédié : <http://www.polytech.univ-nantes.fr/kappalab>. Pour résoudre les programmes linéaires, le *package* LpSolve (Berkelaar et al. 2005) est utilisé. Les programmes quadratiques strictement convexes sont résolus à l'aide du *package* Quadprog (Turlach & Weingessel 2004). Enfin, les programmes quadratiques non nécessairement strictement convexes sont résolus à l'aide de la routine *ipop* du *package* Kernlab (Karatzoglou, Smola, Hornik & Zeileis 2004).

3.4 Exemples d'utilisation des méthodes d'identification

L'objectif de cette dernière section est de donner des exemples d'utilisation des méthodes d'identification présentées dans la section précédente. Les différentes étapes pratiques à mettre en œuvre pour pouvoir réaliser les calculs à l'aide du *package* Kappalab seront également explicitées.

3.4.1 Description du problème

Nous considérons une version étendue du problème présenté dans (Kojadinovic 2006a) concernant l'évaluation d'étudiants dans un institut formant des économétriciens. Les étudiants sont évalués selon 5 matières : statistiques (S), probabilités (P), économie (E), management (M) et anglais (A). Les évaluations partielles sur l'échelle $[0, 20]$ de sept étudiants, a , b , c , d , e , f et g , sont données dans le Tableau 3.1. Elles sont supposées commensurables.

TAB. 3.1 – Évaluations partielles des sept étudiants.

Étudiant	S	P	E	M	A	Moyenne
a	18	11	11	11	18	13.80
b	18	11	18	11	11	13.80
c	11	11	18	11	18	13.80
d	18	18	11	11	11	13.80
e	11	11	18	18	11	13.80
f	11	11	18	11	11	12.40
g	11	11	11	11	18	12.40

Le décideur, occupant le rôle de directeur pédagogique de l'institut, donne un peu plus d'importance aux statistiques et aux probabilités qu'aux autres matières. Il considère également qu'il y a trois groupes de matières : les statistiques et les probabilités, l'économie et le management, et l'anglais, et que les matières au sein des deux premiers groupes évaluent des aptitudes similaires, c.-à-d. elles sont quelque peu substitutives. De plus, si un étudiant est bon en statistiques ou en probabilités (resp. mauvais en statistiques et en probabilités), il est préférable selon le décideur qu'il soit bon en anglais (resp. en économie ou en management) plutôt qu'en économie ou en management (resp. en anglais). Ce raisonnement conduit au rangement suivant des 7 étudiants décrits dans le Tableau 3.1 :

$$a \succ_O b \succ_O c \succ_O d \succ_O e \succ_O f \succ_O g. \quad (3.3)$$

Enfin, le décideur considère que deux étudiants sont significativement différents si et seulement si leurs utilités globales diffèrent d'au moins une demi-unité.

En considérant les étudiants a et b , et f et g , il peut être vérifié que les 5 critères ne sont pas préférentiellement indépendants, ce qui implique qu'il n'existe pas de modèle d'utilité additive qui puisse numériquement représenter le préordre ci-dessus.

Afin de pouvoir utiliser les méthodes d'identification présentées dans la section précédente et implémentées dans Kappalab, nous créons tout d'abord 7 vecteurs R représentant les étudiants :

```
> a <- c(18,11,11,11,18)
```

```
> b <- c(18,11,18,11,11)
> c <- c(11,11,18,11,18)
> d <- c(18,18,11,11,11)
> e <- c(11,11,18,18,11)
> f <- c(11,11,18,11,11)
> g <- c(11,11,11,11,18)
```

Le symbole `>` représente l'invite de commande dans le terminal R, le symbole `<-` désigne l'opérateur d'assignation et `c` est la fonction R permettant la création de vecteurs.

3.4.2 Approches fondées sur les moindres carrés

Pour pouvoir appliquer les méthodes présentées dans la Section 3.3.1 fondées sur les moindres carrés, les 7 vecteurs créés précédemment représentant les étudiants doivent tout d'abord être regroupés dans une matrice 7 lignes, appelée ici `C`, en utilisant la fonction `rbind` (*row bind*) de création de matrices :

```
> C <- rbind(a,b,c,d,e,f,g)
```

Ensuite, le décideur doit fournir les évaluations globales souhaitées pour les sept étudiants. Bien qu'il soit en pratique peu réaliste de considérer que ces informations puissent être facilement fournies, nous supposons dans la suite que le décideur assigne à *a*, *b*, *c*, *d*, *e*, *f* et *g* les notes suivantes : 15, 14.5, 14, 13.5, 13, 12.5 et 12 respectivement. Ces utilités souhaitées sont sauveées dans un vecteur R de 7 éléments :

```
> overall <- c(15,14.5,14,13.5,13,12.5,12)
```

La méthode exacte fondée sur les moindres carrés (donnant une solution optimale mais pas nécessairement unique) peut ensuite être appliquée en tapant

```
> ls <- least.squares.capacity.ident(5,2,C,overall)
```

dans le terminal R. Le premier argument définit le nombre de critères, le second l'ordre de *k*-additivité désiré. Les deux derniers arguments correspondent respectivement à la matrice contenant les utilités partielles et au vecteur contenant les utilités globales souhaitées par le décideur. Les résultats de la routine sont sauveés dans une liste R, appelée ici `ls`, contenant toutes les informations nécessaires à leur analyse.

La solution, une capacité 2-additive donnée sous la forme de sa représentation de Möbius, peut être obtenue en tapant

```
> m <- ls$solution
```

et peut être visualisée en saisissant `m` dans le terminal R :

```
> m
      Mobius.capacity
{ }      0.000000
```

{1}	0.311650
{2}	0.176033
...	...
{3,5}	0.071428
{4,5}	0.001752

Comme cela a été évoqué dans la Section 3.3.1, dans le cas de l'exemple considéré, la solution obtenue n'est probablement pas unique (Miranda & Grabisch 1999).

L'intégrale de Choquet de a par exemple par rapport à la solution peut être obtenue en saisissant :

```
> Choquet.integral(m,a)
[1] 15
```

Afin de pouvoir utiliser la méthode des moindres carrés heuristiques proposée par Grabisch (1995b), il est tout d'abord nécessaire de créer la *capacité initiale* comme évoqué dans la Section 3.3.1. Ici, en l'absence de prérequis sur la forme de l'intégrale de Choquet, nous prenons la capacité uniforme sur l'ensemble des matières :

```
> mu.unif <- as.capacity(uniform.capacity(5))
```

La méthode d'identification peut alors être exécutée en tapant :

```
> hls <- heuristic.ls.capa.ident(5,mu.unif,C,overall,alpha=0.05)
```

Le premier argument définit le nombre de critères, le second la capacité initiale, le troisième la matrice des évaluations partielles, le quatrième le vecteur contenant les utilités désirées et le dernier, la valeur du paramètre contrôlant la descente du gradient.

Les utilités globales calculées à l'aide de l'intégrale de Choquet par rapport aux deux solutions obtenues sont données dans les deux dernières colonnes du tableau ci-dessous. La sixième colonne contient les utilités globales souhaitées par le décideur.

	S	P	E	M	A	Dec	Moy	MC	MCH
a	18	11	11	11	18	15.0	13.8	15.0	15.0
b	18	11	18	11	11	14.5	13.8	14.5	14.5
c	11	11	18	11	18	14.0	13.8	14.0	14.0
d	18	18	11	11	11	13.5	13.8	13.5	13.5
e	11	11	18	18	11	13.0	13.8	13.0	13.0
f	11	11	18	11	11	12.5	12.4	12.5	12.5
g	11	11	11	11	18	12.0	12.4	12.0	12.0

Comme nous pouvons le voir, aussi bien la méthode optimale que la méthode heuristique permettent de retrouver les utilités souhaitées par le décideur. Rappelons quand même que la solution renvoyée par la méthode quadratique optimale est 2-additive alors que celle renvoyée par la méthode heuristique est 5-additive.

Les valeurs de Shapley des solutions obtenues peuvent être calculées à l'aide de la fonction `Shapley.value` et sont données dans le tableau suivant :

	S	P	E	M	A
MC	0.29	0.14	0.21	0.13	0.24
MCH	0.24	0.18	0.20	0.16	0.21

Comme nous pouvions nous y attendre, la valeur de Shapley de la solution obtenue par la méthode heuristique est plus uniforme que celle de la solution renvoyée par la méthode quadratique optimale.

3.4.3 Approches linéaire, du minimum de variance et du minimum de distance

Comme nous l'avons évoqué précédemment, les méthodes fondées sur les moindres carrés ne sont pas bien adaptées à la théorie de l'utilité multiattribut car elles requièrent de l'information qu'un décideur ne peut pas toujours fournir. L'approche fondée sur la programmation linéaire ainsi que les méthodes du minimum de variance et du minimum de distance ne nécessitent qu'un préordre partiel sur l'ensemble des objets disponibles, comme celui donné dans l'Eq. (3.3). Ce préordre est naturellement traduit comme suit :

$$C_{m_\nu}(a) > C_{m_\nu}(b) > C_{m_\nu}(c) > C_{m_\nu}(d) > C_{m_\nu}(e) > C_{m_\nu}(f) > C_{m_\nu}(g).$$

avec $\delta_C = 0.5$ comme seuil d'indifférence.

Pratiquement, le seuil est sauvé dans une variable R

```
> delta.C <- 0.5
```

et le préordre partiel est stocké dans une matrice 6 lignes

```
> Acp <- rbind(c(a,b,delta.C),
               c(b,c,delta.C),
               c(c,d,delta.C),
               c(d,e,delta.C),
               c(e,f,delta.C),
               c(f,g,delta.C))
```

chaque ligne correspondant à une contrainte de la forme $C_{m_\nu}(u(x)) \geq C_{m_\nu}(u(x')) + \delta_C$.

L'approche linéaire est appliquée en tapant :

```
> lp <- lin.prog.capa.ident(5,2,A.Choquet.preorder = Acp)
```

Le premier argument fixe le nombre de critères, le second définit le degré de k -additivité souhaité pour la solution et le dernier correspond au préordre partiel fourni par le décideur. Toutes les informations nécessaires à l'analyse de la solution sont sauvées dans l'objet R `lp`.

L'approche du minimum de variance est exécutée de façon similaire :

```
> mv <- mini.var.capa.ident(5,2,A.Choquet.preorder = Acp)
```

Avant de pouvoir utiliser l'approche du minimum de distance, il est nécessaire de créer la *capacité initiale*. Le décideur n'ayant pas d'a priori, nous choisissons la capacité uniforme sur l'ensemble des matières, qui peut être créée en tapant :

```
> m.mu <- additive.capacity(c(0.2,0.2,0.2,0.2,0.2))
```

La capacité compatible avec les préférences initiales du décideur la plus proche de la capacité uniforme est obtenue en tapant :

```
> md <- mini.dist.capa.ident(m.mu,2,"global.scores",
                             A.Choquet.preorder = Acp)
```

Le deuxième argument définit l'ordre de k -additivité souhaité pour la solution et le troisième indique laquelle des trois distances disponibles doit être utilisée (Kojadinovic 2006*b*). La chaîne de caractères "global.scores" fait référence à la distance donnée dans l'Eq. (3.2).

Les utilités globales calculées par le biais de l'intégrale de Choquet par rapport aux trois solutions 2-additives sont données dans le tableau ci-dessous :

	S	P	E	M	A	Moy	PL	MV	MD
a	18	11	11	11	18	13.8	18.00	15.25	14.95
b	18	11	18	11	11	13.8	17.36	14.75	14.45
c	11	11	18	11	18	13.8	16.73	14.25	13.95
d	18	18	11	11	11	13.8	16.09	13.75	13.45
e	11	11	18	18	11	13.8	15.45	13.25	12.95
f	11	11	18	11	11	12.4	14.82	12.75	12.45
g	11	11	11	11	18	12.4	14.18	12.25	11.95

Notons que, comme nous pouvions nous y attendre, l'approche linéaire conduit aux utilités les plus dispersées, atteignant la valeur maximale qu'une intégrale de Choquet peut atteindre pour les 7 étudiants (c.-à-d. la valeur 18). Notons aussi que pour les approches du minimum de variance et du minimum de distance, les différences entre les utilités globales de deux étudiants consécutifs dans le préordre fourni par le décideur valent exactement δ_C . Ceci est essentiellement dû au fait que, dans le cas de l'exemple traité, l'objectif des deux méthodes est grosso modo de trouver parmi les intégrales de Choquet compatibles avec les préférences initiales du décideur celle qui est la plus proche de la moyenne arithmétique simple.

Les valeurs de Shapley des solutions 2-additives sont :

	S	P	E	M	A
PL	0.45	0.00	0.27	0.05	0.23
MV	0.27	0.16	0.21	0.14	0.22
MD	0.24	0.18	0.20	0.16	0.22

Les trois solutions désignent, comme nous pouvons le voir, les statistiques (S) comme le critère le plus important. Notons que la solution linéaire est très extrême car les importances globales des probabilités (P) et du management (M) sont très faibles et que celle de S est proche de 0.5. Nous voyons ainsi que les importances globales des critères ne sont pas en accord avec l'orientation de l'institut. En effet, les statistiques (S) et les probabilité (P), et l'économie (E) et le management (M), n'ont pas les mêmes importances. Cela est dû au faible nombre d'étudiants rangés par le décideur et peut être corrigé en imposant des contraintes supplémentaires comme nous le verrons dans la section suivante.

Ce dernier tableau donne les représentations de Möbius des trois solutions 2-additives :

	PL	MV	MD
{}	0.00	0.00	0.00
{1}	0.73	0.34	0.33
{2}	0.00	0.18	0.19
{3}	0.55	0.25	0.21
{4}	0.09	0.15	0.16
{5}	0.45	0.18	0.14
{1,2}	0.00	-0.13	-0.17
{1,3}	-0.36	-0.06	-0.05
{1,4}	0.00	-0.04	-0.06
{1,5}	-0.18	0.09	0.10
{2,3}	0.00	0.02	0.03
{2,4}	0.00	0.11	0.15
{2,5}	0.00	-0.03	-0.02
{3,4}	0.00	-0.08	-0.08
{3,5}	-0.18	0.04	0.08
{4,5}	-0.09	-0.01	0.00

3.4.4 Contraintes supplémentaires sur la valeur de Shapley

Nous supposons maintenant, comme nous l'avons évoqué dans la section précédente, qu'après avoir examiné les valeurs de Shapley des solutions 2-additives obtenues précédemment, le décideur requiert que les statistiques (S) et les probabilités (P), et l'économie (E) et le management (M), aient les mêmes importances globales, c.-à-d. $S \sim_N P$ et $E \sim_N M$.

Ces contraintes supplémentaires sont traduites par $-\delta_\phi \leq \phi_{m_\nu}(S) - \phi_{m_\nu}(P) \leq \delta_\phi$ et $-\delta_\phi \leq \phi_{m_\nu}(E) - \phi_{m_\nu}(M) \leq \delta_\phi$, le seuil d'indifférence δ_ϕ étant supposé pris égal à 0.01 par le décideur. Afin d'encoder ces préférences, une variable R représentant le seuil d'indifférence est d'abord créée :

```
> delta.phi <- 0.01
```

Les inégalités évoquées ci-dessus sont ensuite sauvées dans une matrice 4 lignes

```
> Asp <- rbind(c(1,2,-delta.phi),
               c(2,1,-delta.phi),
               c(3,4,-delta.phi),
               c(4,3,-delta.phi))
```

chaque ligne correspondant à une contrainte de la forme $\phi_{m_\nu}(i) - \phi_{m_\nu}(j) \geq c$, $c \in [-1, 1]$.

La routine implémentant l'approche linéaire est ensuite exécutée en tapant

```
> lp2 <- lin.prog.capa.ident(5,2,A.Choquet.preorder = Acp,
                             A.Shapley.preorder = Asp)
```

dans le terminal R. Les routines implémentant les principes du minimum de variance et du minimum de distance sont appelées de façon similaire.

Les valeurs de Shapley des solutions 2-additives sont :

	S	P	E	M	A
PL	0.23	0.23	0.18	0.18	0.18
MV	0.22	0.21	0.18	0.17	0.22
MD	0.22	0.21	0.18	0.17	0.22

Comme prévu, les solutions satisfont les contraintes supplémentaires imposées par le décideur.

Les utilités globales calculées par le biais de l'intégrale de Choquet par rapport aux solutions 2-additives sont données dans le tableau suivant :

	S	P	E	M	A	Moy	PL	MV	MD
a	18	11	11	11	18	13.8	16.03	15.12	14.84
b	18	11	18	11	11	13.8	15.52	14.62	14.34
c	11	11	18	11	18	13.8	15.01	14.12	13.84
d	18	18	11	11	11	13.8	14.50	13.62	13.34
e	11	11	18	18	11	13.8	13.99	13.12	12.84
f	18	18	11	11	11	12.4	13.48	12.62	12.34
g	11	11	18	11	11	12.4	12.97	12.12	11.84

Cette fois, les trois approches conduisent à des utilités globales très proches. Cela s'explique par le fait que le problème est maintenant plus contraint.

Les indices d'interaction des capacités 2-additives obtenues par l'approche linéaire, l'approche du minimum de variance et l'approche du minimum de distance sont donnés dans les trois tableau ci-dessus :

	S	P	E	M	A
S	NA	-0.27	-0.17	0.00	-0.03
P	-0.27	NA	0.00	0.16	-0.04
E	-0.17	0.00	NA	-0.12	-0.06
M	0.00	0.16	-0.12	NA	-0.07
A	-0.03	-0.04	-0.06	-0.07	NA

	S	P	E	M	A
S	NA	-0.21	-0.05	-0.06	0.10
P	-0.21	NA	0.01	0.15	-0.03
E	-0.05	0.01	NA	-0.12	0.05
M	-0.06	0.15	-0.12	NA	-0.01
A	0.10	-0.03	0.05	-0.01	NA

	S	P	E	M	A
S	NA	-0.21	-0.04	-0.07	0.10
P	-0.21	NA	0.03	0.18	-0.01
E	-0.04	0.03	NA	-0.10	0.09
M	-0.07	0.18	-0.10	NA	0.00
A	0.10	-0.01	0.09	0.00	NA

Comme nous pouvons le remarquer, les statistiques (S) interagissent négativement avec presque toutes les matières, ce qui, encore une fois, n'est pas en accord avec l'orientation de

l'institution. En effet, les statistiques (S) devraient plutôt interagir de façon complémentaire avec toutes les matières sauf avec les probabilités (P). Ceci peut être corrigé en imposant des contraintes supplémentaires sur les indices d'interaction comme nous le verrons dans la prochaine section.

3.4.5 Contraintes supplémentaires sur les indices d'interaction

Supposons finalement, qu'afin d'être en accord avec l'orientation de l'institution, le décideur impose que les matières d'un même groupe présentent une interaction substitutive et que deux matières appartenant à des groupes différents présentent une interaction complémentaire. Ces informations préférentielles supplémentaires peuvent être traduites par les contraintes suivantes :

P	E	M	A	
$-1 \leq I_{m_\nu}(SP) \leq -\delta_I$	$\delta_I \leq I_{m_\nu}(SE) \leq 1$	$\delta_I \leq I_{m_\nu}(SM) \leq 1$	$\delta_I \leq I_{m_\nu}(SA) \leq 1$	S
	$\delta_I \leq I_{m_\nu}(PE) \leq 1$	$\delta_I \leq I_{m_\nu}(PM) \leq 1$	$\delta_I \leq I_{m_\nu}(PA) \leq 1$	P
		$-1 \leq I_{m_\nu}(EM) \leq -\delta_I$	$\delta_I \leq I_{m_\nu}(EA) \leq 1$	E
			$\delta_I \leq I_{m_\nu}(MA) \leq 1$	M

où δ_I , supposé fixé à 0.05 par le décideur, est un seuil qui peut être interprété comme la valeur minimale en valeur absolue d'un indice d'interaction pour qu'il soit considéré comme significativement différent de 0.

Afin d'encoder ces informations préférentielles supplémentaires, une variable R représentant le seuil est d'abord créée :

```
> delta.I <- 0.05
```

Les contraintes évoquées ci-dessus sont ensuite sauvées dans la matrice 10 lignes suivante

```
> Aii <- rbind(c(1,2,-1,-delta.I),
               c(1,3,delta.I,1),
               c(1,4,delta.I,1),
               c(1,5,delta.I,1),
               c(2,3,delta.I,1),
               c(2,4,delta.I,1),
               c(2,5,delta.I,1),
               c(3,4,-1,-delta.I),
               c(3,5,delta.I,1),
               c(4,5,delta.I,1))
```

chaque ligne correspondant à une contrainte de la forme $a \leq I_{m_\nu}(ij) \leq b$, $a, b \in [-1, 1]$.

Il n'y a pas de capacités 2-additives compatibles avec ces contraintes supplémentaires. L'ordre de k -additivité est alors incrémenté et l'approche linéaire est appliquée en tapant :

```
> lp3 <- lin.prog.capa.ident(5,3,
                             A.Choquet.preorder = Acp,
                             A.Shapley.preorder = Asp,
                             A.interaction.interval = Aii)
```


Les routines implémentant les approches du minimum de variance et du minimum de distance sont appelées de façon similaire.

Les valeurs de Shapley des trois solutions 3-additives sont :

	S	P	E	M	A
PL	0.23	0.23	0.16	0.16	0.22
MV	0.23	0.22	0.18	0.18	0.20
MD	0.22	0.21	0.18	0.19	0.21

Les indices d'interaction des trois solutions 3-additives sont donnés dans les trois tableaux ci-dessus :

	S	P	E	M	A
S	NA	-0.30	0.05	0.05	0.12
P	-0.30	NA	0.07	0.14	0.05
E	0.05	0.07	NA	-0.24	0.05
M	0.05	0.14	-0.24	NA	0.05
A	0.12	0.05	0.05	0.05	NA

	S	P	E	M	A
S	NA	-0.13	0.05	0.05	0.05
P	-0.13	NA	0.05	0.05	0.05
E	0.05	0.05	NA	-0.05	0.05
M	0.05	0.05	-0.05	NA	0.05
A	0.05	0.05	0.05	0.05	NA

	S	P	E	M	A
S	NA	-0.21	0.05	0.05	0.05
P	-0.21	NA	0.05	0.05	0.05
E	0.05	0.05	NA	-0.12	0.05
M	0.05	0.05	-0.12	NA	0.05
A	0.05	0.05	0.05	0.05	NA

Comme prévu, les contraintes supplémentaires imposées par le décideur sont satisfaites.

Les utilités globales calculées par l'intégrale de Choquet par rapport aux trois solutions 3-additives sont présentées dans le tableau ci-dessus :

	S	P	E	M	A	Moy	PL	MV	MD
a	18	11	11	11	18	13.8	14.06	14.26	14.45
b	18	11	18	11	11	13.8	13.55	13.76	13.95
c	11	11	18	11	18	13.8	13.04	13.26	13.45
d	18	18	11	11	11	13.8	12.53	12.76	12.95
e	11	11	18	18	11	13.8	12.02	12.26	12.45
f	18	18	11	11	11	12.4	11.51	11.76	11.95
g	11	11	18	11	11	12.4	11.00	11.26	11.45

3.4.6 Solution plus simple mais approchée

Supposons enfin que par souci de simplicité, le décideur souhaite absolument disposer d'une solution 2-additive. Nous avons néanmoins vu dans la section précédente qu'il n'existe pas de solution 2-additive permettant de satisfaire les contraintes supplémentaires sur la valeur de Shapley et les indices d'interaction. La méthode des moindres carrés généralisés décrite dans la Section 3.3.4 peut alors être utilisée pour obtenir une solution approchée.

Tout d'abord, le préordre sur les étudiants doit être encodé dans une matrice 6 lignes :

```
> rk.proto <- rbind(c(1,2), c(2,3), c(3,4), c(4,5), c(5,6), c(6,7))
```

Les entiers correspondent aux numéros de ligne des alternatives a, b, c, d, e, f et g dans la matrice C définie dans la Section 3.4.2.

La routine implémentant l'approche des moindres carrés généralisés peut ensuite être appelée en tapant :

```
> gls <- ls.ranking.capa.ident(5, 2, C, rk.proto, 0.5,
                             A.Shapley.preorder = Asp,
                             A.interaction.interval = Aii)
```

Le premier argument fixe le nombre de critères, le second l'ordre de k -additivité désiré pour la solution, le troisième correspond à la matrice contenant les évaluations partielles, le quatrième au préordre sur les objets disponibles et le cinquième argument contient la valeur du seuil δ_y . Les deux derniers arguments correspondent aux matrices encodant les contraintes supplémentaires sur la valeur de Shapley et les indices d'interaction respectivement.

Bien qu'il n'existe pas de capacités 2-additives compatibles avec les contraintes supplémentaires imposées, l'approche des moindres carrés généralisés renvoie une solution n'annulant pas la fonction objectif (comme nous pouvions nous y attendre). Le tableau suivant donne les utilités globales $(C_{m_\nu}(u(x)))_{x \in \mathcal{O}}$ dans la dernière colonne et les utilités $(y(x))_{x \in \mathcal{O}}$ dans l'avant-dernière colonne :

	S	P	E	M	A	Moy	y	MCG
a	18	11	11	11	18	13.8	13.94	13.67
b	18	11	18	11	11	13.8	13.44	13.44
c	11	11	18	11	18	13.8	12.94	12.81
d	18	18	11	11	11	13.8	12.44	12.57
e	11	11	18	18	11	13.8	11.94	11.81
f	18	18	11	11	11	12.4	11.44	11.57
g	11	11	18	11	11	12.4	10.94	11.21

Comme nous pouvons le voir, le préordre imposé par le décideur n'est pas violé mais le seuil minimal δ_y n'est pas toujours respecté (par exemple $C_{m_\nu}(f) - C_{m_\nu}(g) < \delta_y$). La valeur de Shapley de cette solution 2-additive est

S	P	E	M	A
0.23	0.22	0.17	0.16	0.22

et les indices d'interaction

	S	P	E	M	A
S	NA	-0.22	0.05	0.05	0.14
P	-0.22	NA	0.06	0.09	0.06
E	0.05	0.06	NA	-0.08	0.15
M	0.05	0.09	-0.08	NA	0.05
A	0.14	0.06	0.15	0.05	NA

lesquels, comme prévu, vérifient les contraintes supplémentaires imposées par le décideur.

Conclusion et perspectives

Les travaux présentés dans ce mémoire s'articulent, comme nous avons pu le voir, autour de deux axes. Le premier axe porte sur la définition et l'axiomatisation d'indices numériques permettant l'*interprétation* des phénomènes modélisés par des mesures non additives éventuellement généralisées. Le second axe a pour objet l'étude de méthodes permettant d'*identifier* les coefficients de ces objets mathématiques à partir d'informations fournies par un décideur. La complémentarité de ces deux directions de recherche ne fait aucun doute, comme nous l'avons vu dans le Chapitre 3. En effet, une fois un modèle décisionnel identifié, il apparaît nécessaire d'expliquer ses principales caractéristiques afin de le rendre plus transparent pour le décideur. Réciproquement, lors de la phase d'identification, la prise en compte de l'information préférentielle du décideur nécessite sa traduction dans le langage du modèle décisionnel, ce qui fait intervenir les différents indices numériques mentionnés dans le Chapitre 2.

De nombreux aspects des travaux présentés restent bien évidemment à approfondir. Quelques directions de recherche sont discutées ci-après.

Intégrale de Choquet et mesurage conjoint

L'un des aspects les plus importants concerne l'utilisation de l'intégrale de Choquet pour la représentation numérique des préférences en aide multicritère à la décision. Bien qu'une construction rigoureuse ait été proposée par Grabisch & Labreuche (2004) (voir aussi Labreuche & Grabisch 2003, Grabisch 2005*b*, Labreuche & Grabisch 2006*a*), il apparaît nécessaire, comme suggéré par Bouyssou (2006), de disposer de résultats théoriques permettant, dans le contexte du mesurage conjoint appliqué à l'aide multicritère à la décision, de savoir sous quelles conditions une relation binaire sur un produit cartésien (non homogène donc) peut être représentée dans le cadre du modèle transitif décomposable par une intégrale de Choquet. Dans le cas où la fonction d'agrégation est l'intégrale de Sugeno, des résultats remarquables ont été récemment obtenus par Slowiński, Greco & Matarazzo (2002) et Bouyssou & Marchant (2005*a*) (voir aussi Greco, Matarazzo & Slowiński 2004, Bouyssou & Marchant 2005*b*). Le cas de l'intégrale de Choquet a été qualifié par Bouyssou (2006), en substance, de **nettement plus complexe**. Afin de m'en persuader, j'ai récemment commencé à effleurer l'abondante littérature sur le mesurage conjoint. J'ai été frappé par la complexité de la caractérisation du modèle classique d'utilité additive (Gonzales 1996). La caractérisation du modèle transitif décomposable à partir d'axiomes utilisant les *traces marginales* m'a en revanche paru un peu moins ardue (Bouyssou & Pirlot 2004, Greco, Matarazzo & Slowiński 2004). En partant de ces résultats, l'intuition pousse à considérer qu'il suffirait "simplement" de rajouter quelques axiomes supplémentaires pour arriver à caractériser le modèle transitif décomposable fondé sur l'intégrale de Choquet. Ces axiomes s'inspireraient par exemple des axiomes d'annulation utilisés pour caractériser le modèle d'utilité additive, car,

après tout, l'intégrale de Choquet généralise la somme (pondérée). Malheureusement, je crois maintenant comprendre pourquoi cela n'est pas "si simple". En effet, il apparaît rapidement très difficile de formuler des axiomes caractérisant l'intégrale de Choquet et n'utilisant, dans leur expression, que la relation de préférence globale.

Mise en œuvre en aide multicritère à la décision

L'une des difficultés principales pour la mise en œuvre des méthodes présentées dans le Chapitre 3 est l'obtention de fonctions d'utilité commensurables. La démarche pratique proposée par Grabisch & Labreuche (2004) consiste grosso modo à procéder en deux temps : il s'agit tout d'abord d'obtenir des fonctions d'utilité commensurables en appliquant la méthode MACBETH (voir p. ex. Labreuche & Grabisch 2003), puis d'utiliser une des méthodes d'identification proposées dans le Chapitre 3 pour déterminer les coefficients de la capacité. D'un point de vue pratique, cette façon de procéder est très coûteuse en temps et demande énormément d'informations au décideur pour la détermination des fonctions d'utilité. Une direction de recherche intéressante serait d'étudier la possibilité de déterminer dans un même temps les fonctions d'utilité et les coefficients de la capacité. Une première approche de ce type a été proposée par Angilella, Greco, Lamantia & Matarazzo (2004) mais est très lourde à mettre œuvre et semble conduire à des solutions assez extrêmes.

Modèles décisionnels fondés sur l'intégrale de Choquet bipolaire

Les indices numériques les plus importants permettant d'interpréter les capacités ont été récemment étendus aux bi-capacités. Il reste cependant à s'assurer que les interprétations de ces indices sont suffisamment naturelles pour qu'ils puissent être utilisés en pratique. En ce qui concerne l'identification de bi-capacités, il semblerait que la plupart des méthodes présentées dans le Chapitre 3 puissent être étendues au cas bipolaire. Ces généralisations ne sont néanmoins pas nécessairement immédiates comme évoqué dans (Kojadinovic 2006*b*, Section 3).

Loi de l'intégrale de Choquet

La loi exacte de l'intégrale de Choquet dans le cas uniforme a été déterminée par Marichal & Kojadinovic (2006) (cf. Section 2.5). Un sujet d'étude intéressant serait d'essayer d'obtenir des résultats du même type pour d'autres distributions des intégrandes. Grabisch & Raufaste (2006) se sont par exemple intéressés empiriquement à ce problème dans le cas d'intégrandes normalement distribuées. Les résultats présentés dans (Govindarajulu 1968, Matsunawa 1985) pour les combinaisons linéaires de statistiques d'ordre suggèrent qu'il est peut-être possible, au moins asymptotiquement, de déterminer la loi de l'intégrale de Choquet dans ce cas.

Axiomatisation des mesures d'uniformité de capacités

Une autre direction de recherche possible concerne l'étude et l'axiomatisation des mesures d'uniformité de capacités normalisées. En effet, les résultats obtenus jusqu'à présent (Kojadinovic, Marichal & Roubens 2005, Honda & Grabisch 2006) reposent sur des axiomes qui ne traduisent pas l'essence même de la notion d'uniformité. En ce sens, il serait intéressant d'essayer de généraliser l'approche adoptée par Morales, Pardo & Vajda (1996) pour les entropies probabilistes

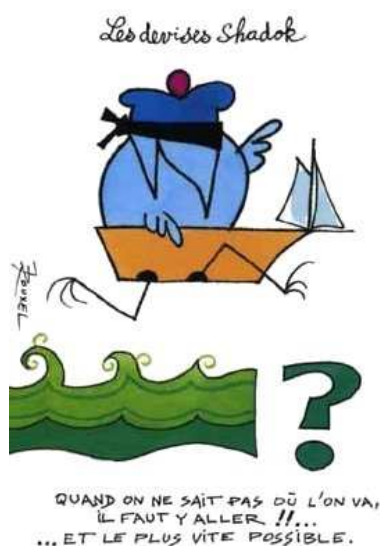
qui lie la notion d'uniformité à celle de *Schur-concavité*, elle-même liée au concept de *dominance de Lorenz* (voir p. ex. Marshall & Olkin 1979, Dubois & Hüllermeier 2006).

Approfondir les liens avec les autres théories

Tout au long de cette synthèse, nous avons vu que les mêmes concepts mathématiques se retrouvent dans la plupart des théories fondées sur les mesures non additives. L'exemple le plus flagrant semble être la valeur de Shapley. Il serait ainsi intéressant d'essayer de donner, dans d'autres contextes théoriques, des interprétations convaincantes de certaines notions définies initialement dans le contexte de l'aide multicritère à la décision fondée sur les intégrales non additives, comme les indices d'interaction et les mesures d'uniformité.

“C'est tout pour aujourd'hui”

Nous terminerons ce mémoire comme nous l'avons commencé. En effet, tu te doutes bien, cher lecteur, que les Shadoks dans leur infinie sagesse ont une devise adaptée à toutes les situations :



Jacques Rouxel, *Les Shadoks*.

Bibliographie

- Aczél, J. & Daróczy, Z. (1975), *On measures of information and their characterizations*, Academic Press, New York–San Francisco–London.
- Agarwal, G. G., Dalpatadu, R. J. & Singh, A. K. (2002), ‘Linear functions of uniform order statistics and B-splines’, *Commun. Stat., Theory Methods* **31**(2), 181–192.
- Ali, M. M. (1973), ‘Content of the frustum of a simplex’, *Pacific Journal of Mathematics* **48**(2), 313–322.
- Angilella, S., Greco, S., Lamantia, F. & Matarazzo, B. (2004), ‘Assessing non-additive utility for multicriteria decision aid’, *European Journal of Operational Research* **158**, 734–744.
- Aubin, J. (1981), ‘Cooperative fuzzy games’, *Mathematics of Operations Research* **6**, 1–13.
- Bana e Costa, C. & Vansnick, J. (1999), Preference relations in MCDM, in T. Gal & T. H. T. Steward, eds, ‘MultiCriteria Decision Making : Advances in MCDM models, algorithms, theory and applications’, Kluwer.
- Bana e Costa, C., Corte, J. D. & Vansnick, J. (2005), On the mathematical foundations of MAC-BETH, in J. Figueira, S. Greco & M. Ehrgott, eds, ‘Multiple Criteria Decision Analysis : State of the Art Surveys’, Springer, pp. 409–442.
- Banzhaf, J. (1965), ‘Weighted voting doesn’t work : A mathematical analysis’, *Rutgers Law Review* **19**, 317–343.
- Ben-Or, M. & Linial, N. (1990), Collective coin flipping, in ‘Randomness and Computation’, Academic Press, New York, pp. 91–115. (Earlier version : Collective coin flipping, robust voting games and minima of Banzhaf values, Proc. 26th IEEE Symposium on the Foundation of Computer Sciences, Portland, 1985, pp. 408–416).
- Benvenuti, P., Mesiar, R. & Vivona, D. (2002), Monotone set functions-based integrals, in E. Pap, ed., ‘Handbook of Measure Theory’, Elsevier Science, pp. 1329–1379.
- Berkelaar, M. et al. (2005), **lpSolve** : *Interface to LPSolve v.5 to solve linear/integer programs*. R package version 1.1.9.
- Bilbao, J. (1998), ‘Axioms for the Shapley value on convex geometries’, *European Journal of Operational Research* **110**, 365–372.
- Bilbao, J. (2000), *Cooperative games on combinatorial structures*, Kluwer Academic Publishers.
- Bilbao, J., Fernández, J., Jiménez, N. & López, J. (2006), ‘The value for bicooperative games’, *Annals of Operations Research* p. in press.
- Bourgain, J., Kahn, J., Kalai, G., Katznelson, Y. & Linial, N. (1992), ‘The influence of variables in product spaces’, *Isr. J. Math.* **77**(1-2), 55–64.
- Bouyssou, D. (2006), A conjoint measurement view on fuzzy integrals, in ‘27th Linz Seminar on Fuzzy Set Theory "Preferences, Games and Decisions"', Linz, Austria, pp. 44–45.

- Bouyssou, D. & Marchant, T. (2005a), 'An axiomatic approach to noncompensatory sorting methods in MCDM I : The case of two categories', *European Journal of Operational Research* p. in press.
- Bouyssou, D. & Marchant, T. (2005b), 'An axiomatic approach to noncompensatory sorting methods in MCDM II : More than two categories', *European Journal of Operational Research* p. in press.
- Bouyssou, D. & Pirlot, M. (2004), 'Preferences for multi-attributed alternatives : Traces, dominance, and numerical representations', *J. of Mathematical Psychology* **48**, 167–185.
- Bouyssou, D. & Pirlot, M. (2005), Conjoint measurement tools for MCDM : A brief introduction, in J. Figuera, S. Greco & M. Ehrgott, eds, 'Multiple Criteria Decision Analysis : State of the Art Surveys', Springer, pp. 73–132.
- Bouyssou, D. & Pirlot, M. (2006), Mesurage conjoint et modèles relationnels de préférences, in D. Bouyssou, D. Dubois, M. Pirlot & H. Prade, eds, 'Concepts et Méthodes pour l'Aide à la Décision', Vol. 3 of *IC2*, Hermès, pp. 63–120.
- Chateauneuf, A. (1994), 'Modeling attitudes towards uncertainty and risk through the use of Choquet integral', *Annals of Operations Research* **52**, 3–20.
- Chateauneuf, A. & Jaffray, J.-Y. (1989), 'Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion', *Mathematical Social Sciences* **17**(3), 263–283.
- Choquet, G. (1953), 'Theory of capacities', *Annales de l'Institut Fourier* **5**, 131–295.
- Cover, T. & Thomas, J. (1991), *Elements of Information Theory*, John Wiley and Sons.
- Curiel, I. (1997), *Cooperative Game Theory and Applications*, Kluwer Academic.
- Davis, P. J. (1975), *Interpolation and approximation. 2nd ed.*, Dover Books on Advanced Mathematics. New York : Dover Publications.
- de Boor, C. (1972), 'On calculating with *b*-splines', *Journal of Approximation Theory* **6**, 50–62.
- Dempster, A. (1967), 'Upper and lower probabilities induced by a multivalued mapping', *Ann. Math. Statist.* **38**, 325–339.
- Denneberg, D. (1994), *Non-additive Measure and Integral*, Kluwer Academic, Dordrecht.
- Denneberg, D. (2000), Non-additive measure and integral, basic concepts and their role for applications, in M. Grabisch, T. Murofushi & M. Sugeno, eds, 'Fuzzy Measures and Integrals : Theory and Applications', Physica-Verlag, pp. 42–69.
- Dubois, D. & Hüllermeier, E. (2006), 'Comparing probability measures using possibility theory : A notion of relative peakedness'. working paper.
- Dubois, D. & Prade, H. (1987), 'Properties of measures of information in evidence and possibility theories', *Fuzzy Sets and Systems* **24**, 161–182.
- Dubois, D. & Prade, H. (1988), *Possibility theory : An approach to the computerized processing of uncertainty*, Plenum Press.
- Dubois, D. & Ramer, A. (1993), 'Extremal properties of belief measures in the theory of evidence', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **1**(1), 57–68.
- Dukhovny, A. (2002), 'General entropy of general measures', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **10**(3), 213–226.
- Ebanks, B., Sahoo, P. & Sander, W. (1997), *Characterizations of information measures*, World Scientific, Singapore.

-
- Esteban, M. & Morales, D. (1995), 'A summary on entropy statistics', *Kybernetika* **31**(4), 337–346.
- Faddeev, D. (1957), 'Zum begriff der entropie einer endlichen wahrscheinlichkeitschemes', *Arbeit zur Informationstheorie, Deutscher Verlag der Wissenschaften*.
- Felsenthal, D. & Machover, M. (1997), 'Ternary voting games', *Int. J. of Game Theory* **26**, 335–351.
- Fujimoto, K. & Murofushi, T. (2000), Hierarchical decomposition of the Choquet integral, in M. Grabisch, T. Murofushi & M. Sugeno, eds, 'Fuzzy Measures and Integrals : Theory and Applications', Physica-Verlag, pp. 95–103.
- Fujimoto, K. & Murofushi, T. (2005), 'Some characterizations of k -monotonicity through the bipolar Möbius transform in bi-capacities', *Journal of Advanced Computational Intelligence and Intelligent Informatics* **9**(5), 484–495.
- Fujimoto, K., Kojadinovic, I. & Marichal, J.-L. (2006), 'Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices', *Games and Economic Behavior* **55**, 72–99.
- Gonzales, C. (1996), Utilités additives : existence et construction, PhD thesis, Université Paris 6, France.
- Govindarajulu, Z. (1968), 'Asymptotic normality of linear combinations of functions of order statistics', *Proc. Natl. Acad. Sci. USA* **59**(3), 713–719.
- Grabisch, M. (1992), 'The application of fuzzy integrals in multicriteria decision making', *European Journal of Operational Research* **89**, 445–456.
- Grabisch, M. (1995a), 'Fuzzy integral in multicriteria decision making', *Fuzzy Sets and Systems* **69**, 279–298.
- Grabisch, M. (1995b), A new algorithm for identifying fuzzy measures and its application to pattern recognition, in 'Int. 4th IEEE Conf. on Fuzzy Systems', Yokohama, Japan, pp. 145–150.
- Grabisch, M. (1996), 'The representation of importance and interaction of features by fuzzy measures', *Pattern Recognition Letters* **17**(6), 567–575.
- Grabisch, M. (1997a), 'Alternative representations of discrete fuzzy measures for decision making', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **5**(5), 587–607.
- Grabisch, M. (1997b), ' k -order additive discrete fuzzy measures and their representation', *Fuzzy Sets and Systems* **92**(2), 167–189.
- Grabisch, M. (2000), The interaction and Möbius representations of fuzzy measures on finite spaces, k -additive measures : A survey, in M. Grabisch, T. Murofushi & M. Sugeno, eds, 'Fuzzy Measures and Integrals : Theory and Applications', Physica-Verlag, pp. 71–93.
- Grabisch, M. (2004), The Choquet integral as a linear interpolator, in '10th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)', Perugia, Italy, pp. 373–378.
- Grabisch, M. (2005a), 'Capacities and games on lattices : A survey of results', *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems* p. in press.
- Grabisch, M. (2005b), Évaluation subjective, in D. Bouyssou, D. Dubois, M. Pirlot & H. Prade, eds, 'Concepts et Méthodes pour l'Aide à la Décision', IC2, Hermès, p. in press.
- Grabisch, M. (2005c), Une approche constructive de la décision multicritère, working paper.

- Grabisch, M. & Labreuche, C. (2002), 'The symmetric and asymmetric Choquet integrals on finite spaces for decision making', *Statistical Papers* **43**, 37–52.
- Grabisch, M. & Labreuche, C. (2004), Fuzzy measures and integrals in MCDA, in J. Figueira, S. Greco & M. Ehrgott, eds, 'Multiple Criteria Decision Analysis', Springer, pp. 563–608.
- Grabisch, M. & Labreuche, C. (2005a), 'Bi-capacities I : Definition, Möbius transform and interaction', *Fuzzy Sets and Systems* **151**, 211–236.
- Grabisch, M. & Labreuche, C. (2005b), 'Bi-capacities II : The Choquet integral', *Fuzzy Sets and Systems* **151**, 237–259.
- Grabisch, M. & Labreuche, C. (2005c), Bi-capacities : Towards a generalization of cumulative prospect theory, working paper.
- Grabisch, M. & Labreuche, C. (2005d), A general construction for unipolar and bipolar interpolative aggregation, in '4th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2005)', Barcelona, Spain, pp. 916–921.
- Grabisch, M. & Nicolas, J.-M. (1994), 'Classification by fuzzy integrals : Performance and test', *Fuzzy Sets and Systems* **65**, 255–271.
- Grabisch, M. & Raufaste, E. (2006), 'An experimental study of statistical properties of Choquet and Sugeno integrals'. working paper.
- Grabisch, M. & Roubens, M. (1999), 'An axiomatic approach to the concept of interaction among players in cooperative games', *Internat. J. Game Theory* **28**(4), 547–565.
- Grabisch, M., Kojadinovic, I. & Meyer, P. (2005), *kappalab : Non additive measure and integral manipulation functions*. R package version 0.2.
- Grabisch, M., Kojadinovic, I. & Meyer, P. (2006), 'A review of capacity identification methods for Choquet integral based multi-attribute utility theory : Applications of the Kappalab R package'. submitted.
- Grabisch, M., Labreuche, C. & Vansnick, J. (2003), 'On the extension of pseudo-Boolean functions for the aggregation of interacting criteria', *European Journal of Operational Research* **148**, 28–47.
- Grabisch, M., Marichal, J.-L. & Roubens, M. (2000), 'Equivalent representations of set functions', *Math. Oper. Res.* **25**(2), 157–178.
- Grabisch, M., Nguyen, H. & Walker, E. (1995), *Fundamentals of uncertainty calculi with applications to fuzzy inference*, Kluwer Academic, Dordrecht.
- Greco, S., Matarazzo, B. & Slowinski, R. (2002), Bipolar Sugeno and Choquet integrals, in 'Seventh Meeting of the EURO Working Group on Fuzzy Sets (EUROFUSE)', Varenna, Italy, pp. 191–196.
- Greco, S., Matarazzo, B. & Slowinski, R. (2003), The axiomatic basis of multicriteria noncompensatory preferences, in 'Forth International Workshop on Preferences and Decisions', Trento, Italy, pp. 81–86.
- Greco, S., Matarazzo, B. & Slowiński, R. (2004), 'Axiomatic characterization of a general utility function in terms of conjoint measurement and rough-set decision rules', *European Journal of Operational Research* **158**(2), 271–292.
- Hammer, P. & Rudeanu, S. (1968), *Boolean methods in operation research and related areas*, Springer, Berlin.
- Harsanyi, J. C. (1963), 'A simplified bargaining model for the n-person cooperative game', *International Economic Review* **4**, 194–220.

-
- Havrda, J. & Charvat, F. (1967), ‘Quantification method in classification processes : Concept of structural α -entropy’, *Kybernetika* **3**, 30–35.
- Honda, A. & Grabisch, M. (2006), ‘An axiomatization of entropy of capacities on set systems’. working paper.
- Ishii, K. & Sugeno, M. (1996), ‘A model of human evaluation process using fuzzy measure’, *Int. J. Man-Machine Studies* **67**, 242–257.
- Jaynes, E. (2003), *Probability Theory : The Logic of Science*, Cambridge University Press.
- Joe, H. (1989), ‘Relative entropy measures of multivariate dependence’, *J. Am. Statist. Assoc.* **84**, 157–164.
- Kahn, J., Kalai, G. & Linial, N. (1988), The influence of variables on Boolean functions, in ‘Proc. 29th Annual Symposium on Foundations of Computational Science’, Computer Society Press, pp. 68–80.
- Karatzoglou, A., Smola, A., Hornik, K. & Zeileis, A. (2004), ‘**kernlab** : An S4 package for kernel methods in R’, *Journal of Statistical Software* **11**(9), 1–20.
- Keeney, R. L. & Raiffa, H. (1976), *Decision with multiple objectives*, Wiley, New-York.
- Khinchin, A. (1957), *Mathematical foundations of information theory*, Dover.
- Klir, G. & Yuan, B. (1995), *Fuzzy Sets and Fuzzy Systems : Theory and Applications*, Prentice Hall.
- Klir, G. J. & Smith, R. M. (2001), ‘On measuring uncertainty and uncertainty-based information : Recent developments’, *Annals of Mathematics and Artificial Intelligence* **32**, 5–33.
- Klir, G. J. & Wierman, M. J. (1999), *Uncertainty-based information : Elements of generalized information theory*, Physica-Verlag. Second edition.
- Kojadinovic, I. (2005), ‘An axiomatic approach to the measurement of the amount of interaction among criteria or players’, *Fuzzy Sets and Systems* **152**, 417–435.
- Kojadinovic, I. (2006a), ‘Minimum variance capacity identification’, *European Journal of Operational Research* p. in press.
- Kojadinovic, I. (2006b), ‘Quadratic distances for capacity and bi-capacity approximation and identification’, *Quarterly Journal of Operations Research* p. in press.
- Kojadinovic, I. (2006c), ‘A weight-based approach to the measurement of the interaction among criteria in the framework of aggregation by the bipolar choquet integral’, *European Journal of Operational Research* p. in press.
- Kojadinovic, I. & Marichal, J.-L. (2006), ‘Entropy of bi-capacities’, *European Journal of Operational Research* p. in press.
- Kojadinovic, I., Marichal, J.-L. & Roubens, M. (2005), ‘An axiomatic approach to the definition of the entropy of a discrete Choquet capacity’, *Information Sciences* **172**, 131–153.
- Krantz, D., Luce, R., Suppes, P. & Tversky, A. (1971), *Foundations of measurement, volume 1 : Additive and polynomial representations*, Academic Press.
- Kullback, S. & Leibler, R. A. (1951), ‘On information and sufficiency’, *Ann. Math. Stat.* **22**, 79–86.
- Labreuche, C. & Grabisch, M. (2003), ‘The Choquet integral for the aggregation of interval scales in multicriteria decision making’, *Fuzzy Sets and Systems* **137**, 11–16.
- Labreuche, C. & Grabisch, M. (2006a), ‘Generalized Choquet-like aggregation functions for handling bipolar scales’, *European Journal of Operational Research* **172**(3), 931–955.

- Labreuche, C. & Grabisch, M. (2006*b*), 'A value for bi-cooperative games'. working paper.
- Lange, F. & Grabisch, M. (2005), New axiomatizations of the Shapley interaction index for bi-capacities, *in* '4th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2005)', Barcelona, Spain, pp. 198–203.
- Lovász, L. (1983), Submodular function and convexity, *in* A. Bachem, M. Grötschel & B. Korte, eds, 'Mathematical programming : The state of the art', Springer-Verlag, pp. 235–257.
- Marichal, J.-L. (1998), Aggregation operators for multicriteria decision aid, PhD thesis, University of Liège, Liège, Belgium.
- Marichal, J.-L. (2000*a*), 'An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria', *IEEE Transactions on Fuzzy Systems* **8**(6), 800–807.
- Marichal, J.-L. (2000*b*), Behavioral analysis of aggregation in multicriteria decision aid, *in* J. Fodor, B. D. Baets & P. Perny, eds, 'Preferences and Decisions under Incomplete Knowledge', Physica-Verlag, pp. 153–178.
- Marichal, J.-L. (2000*c*), 'The influence of variables on pseudo-boolean functions with applications to game theory and multicriteria decision making', *Discrete Applied Mathematics* **107**, 139–164.
- Marichal, J.-L. (2002*a*), Aggregation of interaction criteria by means of the discrete Choquet integral, *in* 'Aggregation operators : New trends and applications', Vol. 97 of *Studies in Fuzziness and Soft Computing*, Physica Verlag, pp. 224–244.
- Marichal, J.-L. (2002*b*), 'Entropy of discrete Choquet capacities', *European Journal of Operational Research* **3**(137), 612–624.
- Marichal, J.-L. (2004*a*), k -intolerant capacities and Choquet integrals, *in* '10th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)', Perugia, Italia, pp. 601–608.
- Marichal, J.-L. (2004*b*), 'Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral', *European Journal of Operational Research* **155**(3), 771–791.
- Marichal, J.-L. (2005), 'Cumulative distribution functions and moments of lattice polynomials', *Statistics and Probability Letters* p. in press.
- Marichal, J.-L. & Kojadinovic, I. (2006), The distribution of linear combinations of lattice polynomials from the uniform distribution, working paper.
- Marichal, J.-L. & Roubens, M. (2000*a*), 'Determination of weights of interacting criteria from a reference set', *European Journal of Operational Research* **124**, 641–650.
- Marichal, J.-L. & Roubens, M. (2000*b*), 'Entropy of discrete fuzzy measures', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **8**(6), 625–640.
- Marichal, J.-L., Kojadinovic, I. & Fujimoto, K. (2006), 'Axiomatic characterizations of generalized values', *Discrete Applied Mathematics* p. in press.
- Marichal, J.-L., Meyer, P. & Roubens, M. (2005), 'Sorting multiattribute alternatives : The TOMASO method', *Computers & Operations Research* **32**(4), 861–877.
- Marshall, A. & Olkin, I. (1979), *Inequalities : Theory of majorization and its applications*, Academic Press, New-York.
- Matsunawa, T. (1985), 'The exact and approximate distributions of linear combinations of selected order statistics from a uniform distribution', *Ann. Inst. Stat. Math.* **37**, 1–16.

-
- Meyer, P. & Roubens, M. (2005), Choice, ranking and sorting in fuzzy Multiple Criteria Decision Aid, in J. Figueira, S. Greco & M. Ehrgott, eds, 'Multiple Criteria Decision Analysis : State of the Art Surveys', Springer, New York, pp. 471–506.
- Miranda, P. & Grabisch, M. (1999), 'Optimization issues for fuzzy measures', *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems* **7**, 545–560.
- Morales, D., Pardo, L. & Vajda, I. (1996), 'Uncertainty of discrete stochastic systems : General theory and statistical theory', *IEEE Trans. on System, Man and Cybernetics* **26**(11), 1–17.
- Mori, T. & Murofushi, T. (1989), An analysis of evaluation model using fuzzy measure and the Choquet integral, in '5th Fuzzy System Symposium', Kobe, Japan, pp. 207–212. In Japanese.
- Moulin, H. (2004), *Fair Division and Collective Welfare*, MIT Press.
- Murofushi, T. & Soneda, S. (1993), Techniques for reading fuzzy measures (iii) : Interaction index, in '9th Fuzzy System Symposium', Sapporo, Japan, pp. 693–696.
- Murofushi, T. & Sugeno, M. (2000), Fuzzy measures and fuzzy integrals, in M. Grabisch, T. Murofushi & M. Sugeno, eds, 'Fuzzy Measures and Integrals : Theory and Applications', Physica-Verlag, pp. 3–41.
- Narukawa, Y. & Murofushi, T. (2004), Decision modelling using the Choquet integral, in 'Modeling Decisions for Artificial Intelligence (MDAI 2004)', Lecture Notes in Artificial Intelligence (LNAI 3131), Springer-Verlag, Barcelona, Spain, pp. 183–193.
- Peleg, B. & Sudhölter, P. (2003), *Introduction to the theory of cooperative games*, Kluwer Academic Publisher.
- Powell, M. (1981), *Approximation theory and methods*, Cambridge University Press.
- R Development Core Team (2005), *R : A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3.
- Rota, G.-C. (1964), 'On the foundations of combinatorial theory. I. Theory of Möbius functions', *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **2**, 340–368 (1964).
- Roubens, M. (1996), Interaction between criteria and definition of weights in MCDA problems, in '44th meeting of the European Working Group "Multicriteria Aid for Decisions', Brussels, Belgium.
- Rudin, W. (1975), *Analyse réelle et complexe*, Masson, Paris.
- Schmeidler, D. (1989), 'Subjective probability and expected utility without additivity', *Econometrica* **57**(3), 571–587.
- Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton Univ. Press.
- Shannon, C. E. (1948), 'A mathematical theory of communication', *Bell Systems Technical Journal* **27**, 379–623.
- Shapley, L. S. (1953), A value for n -person games, in 'Contributions to the theory of games, vol. 2', Annals of Mathematics Studies, no. 28, Princeton University Press, Princeton, N. J., pp. 307–317.
- Shapley, L. S. & Shubik, M. (1954), 'A method for evaluating the distribution of power in a committee system', *American Political Science Review* **48**, 787–792.
- Slowiński, R., Greco, S. & Matarazzo, B. (2002), 'Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification', *Control and Cybernetics* **31**(4), 1005–1035.

- Smets, P. (2005), ‘Decision making in the TBM : The necessity of the pignistic transformation’, *International Journal of Approximate Reasoning* **38**, 133–147.
- Smets, P. & Kennes, R. (1994), ‘The Transferable Belief Model’, *International Journal of Approximate Reasoning* **66**, 191–243.
- Sugeno, M. (1974), Theory of fuzzy integrals and its applications, PhD thesis, Tokyo Institute of Technology, Tokyo, Japan.
- Turlach, B. & Weingessel, A. (2004), *quadprog : Functions to solve quadratic programming problems*. R package version 1.4-7.
- Tversky, A. & Kahneman, D. (1992), ‘Advances in prospect theory : Cumulative representation of uncertainty’, *Journal of Risk and Uncertainty* **5**, 297–323.
- Varsi, G. (1973), ‘The multidimensional content of the frustrum of the simplex’, *Pacific Journal of Mathematics* **46**(1), 303–314.
- Vincke, P. (1992), *Multicriteria Decision-aid*, Wiley.
- Vitali, G. (1925), ‘Sulla definizione di integrale delle funzioni di una variabile’, *Ann. Mat. Pura ed Appl.* **IV**(2), 111–121.
- Šipoš, J. (1979), ‘Integral with respect to a pre-measure’, *Math. Slovaca* **29**, 141–155.
- Wakker, P. (1989), *Additive Representations of Preferences : A New Foundation of Decision Analysis*, Kluwer Academic Publishers, Dordrecht.
- Yager, R. D. (1988), ‘On ordered weighted averaging aggregation operators in multicriteria decision making’, *IEEE Trans. on Systems, Man and Cybernetics* **18**, 183–190.
- Yager, R. D. (2000), ‘On the entropy of fuzzy measures’, *IEEE Transactions on Fuzzy Systems* **8**(4), 453–461.

Annexes

A

Articles de revues joints

Afin de permettre l'approfondissement de certains des concepts évoqués dans ce mémoire, les quatre articles suivants sont inclus ci-après dans leur intégralité :

- K. Fujimoto, I. Kojadinovic and J-L. Marichal (2006), Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices, *Games and Economic Behavior* 55 : 1, pages 72-99, Elsevier.
- I. Kojadinovic (2006), A weight-based approach to the measurement of the interaction among criteria in the framework of aggregation by the bipolar Choquet integral, *European Journal of Operational Research*, à paraître, disponible sur www.sciencedirect.com, Elsevier.
- I. Kojadinovic (2006), Quadratic distances for capacity and bi-capacity approximation and identification, *A Quarterly Journal of Operations Research (4 OR)*, disponible sur www.springerlink.com, Springer.
- I. Kojadinovic (2005), Relevance measures for subset variable selection in regression problems based on k-additive mutual information, *Computational Statistics and Data Analysis* 49, pages 1205-1227, Elsevier.

Le dernier article traite de l'application de la notion de k -additivité à la sélection de variables dans le contexte de la modélisation statistique, aspect qui n'a pas été abordé dans les trois premiers chapitres de cette synthèse.

Les autres articles dont je suis (co-)auteur sont disponibles en ligne, sous forme de versions préliminaires, à l'adresse :

<http://www.sciences.univ-nantes.fr/info/perso/permanents/kojadinovic/>